

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw.model

Run on: September 28, 2002, 17:53:32 ; Search time 62.3 seconds
(without alignments)
1201.665 Million cell updates/sec

Title: US-09-899-471-2
Perfect score: 3605
Sequence: 1 MPVSWFLSLALGNRPVVS.....SSEAPGCEEMDGFCTTLE 674

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
A.Geneseq_032802:*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2352.5	65.3	692	22	AA61880	Human cytokine rec
2	2339	64.9	705	22	AA029322	Human PRO polypept
3	2339	64.9	705	22	AA004956	Human Interleukin
4	2339	64.9	705	22	AA087606	Human PRO20040. H
5	2339	64.9	705	22	AA61884	Chimeric Zcytor14
6	2277	63.2	675	22	AA61885	Chimeric Zcytor14
7	2263.5	62.8	688	22	AA61883	Chimeric Zcytor14
8	1851.5	51.4	575	22	AA61881	Human variant Zcyt
9	956.5	26.6	309	21	AA176143	Human secreted pro
10	841.5	23.3	332	22	AA025864	Human protein sequ
11	838.5	23.3	267	22	AA088448	Human membrane or

12	591.5	16.4	204	22	AAE10920	Human gene 12 enco
13	310.5	8.6	617	21	AAV76048	Murine skin cell p
14	310.5	8.6	617	22	AA655987	Skin cell protein,
15	310	8.6	667	22	AA004957	Human Interleukin
16	167	4.6	667	17	AA004184	Murine Interleukin
17	167	4.6	864	19	AA061271	Mouse Interleukin-
18	167	4.6	864	20	AA092408	Murine IL-17R prot
19	167	4.6	864	21	AAV99935	Murine IL-17R prot
20	167	4.6	864	21	AAV97130	Murine Interleukin
21	167	4.6	864	21	AAV97180	Murine Interleukin
22	167	4.6	864	21	AA03806	Murine IL-17R poly
23	167	4.6	864	22	AA062060	Murine Interleukin
24	167	4.6	864	22	AAV72748	Murine Interleukin
25	151.5	4.2	866	17	AA004185	Human Interleukin-
26	151.5	4.2	866	19	AA061272	Human Interleukin-
27	151.5	4.2	866	20	AA092409	Human IL-17R prote
28	151.5	4.2	866	21	AAV99941	Human IL-17R prote
29	151.5	4.2	866	21	AAV97131	Human Interleukin-
30	151.5	4.2	866	21	AAV97181	Human Interleukin-
31	151.5	4.2	866	21	AA03807	Human Interleukin-
32	151.5	4.2	866	22	AA062066	Human IL-17R (hCTL
33	151.5	4.2	866	22	AAV72754	Human Interleukin
34	149	4.1	330	21	AAV75947	Murine skin cell s
35	149	4.1	330	22	AA055886	Skin cell protein,
36	125	3.5	1042	22	ABG26706	Novel human diagno
37	118.5	3.3	877	22	ABG28379	Novel human diagno
38	114.5	3.2	1013	22	AA081163	Mycobacterium tube
39	114.5	3.2	1711	19	AAW70506	Osteoclasticular pr
40	114.5	3.2	1711	19	AAW70507	Mutant osteoclast
41	108	3.0	752	22	ABB10333	Human CDNA SEQ ID
42	108	3.0	997	22	ABG08088	Novel human diagno
43	105	2.9	1767	22	ABB62726	Drosophila melanog
44	104.5	2.9	1149	22	AAV97640	Apat-1X(Delta12-10
45	104.5	2.9	1237	22	AAV97638	Apat-1MD protein s

ALIGNMENTS

RESULT 1	AA61880	standard; Protein; 692 AA.
ID	AA61880	
XX	AA61880;	
AC	08-MAY-2001. (first entry)	
DT	XX	
XX	XX	
DE	Human cytokine receptor Zcytor14.	
XX	XX	
KW	Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;	
KW	antlinflammatory; gene therapy; vaccine.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
PN	WO200104304-A1.	
XX	XX	
PD	18-JAN-2001.	
XX	XX	
PF	30-JUN-2000; 2000WO-US18383.	
XX	XX	
PR	07-JUL-1999; 99US-0348854.	
XX	XX	
PA	(ZIMO) ZYMOGENETICS INC.	
XX	XX	
PI	Presnell SR, Burkhead SK, Powderder SL;	
XX	XX	
DR	WPI: 2001-112618/12.	
XX	XX	
DR	N-PSDB; AAC85027.	
XX	XX	
PT	New polypeptide encoding a human cytokine receptor Zcytor14, for	
PT	treating inflammation e.g. rheumatoid arthritis	
XX	XX	
PS	Claim 2; Page 2; 112pp; English.	

Sequence 692 AA;

Matches 457; Conservative 66; Mismatches 136;

660 lqpaldsyfthppgtlpapq 677

ID AAU29322 standard; Protein; 705 AA.
 XX AAU29322;
 AC
 DT 18-DEC-2001 (first entry)
 XX
 DE Human PRO polypeptide sequence #299.
 XX
 KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
 KM dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon
 KM adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder
 XX
 OS Homo sapiens.
 XS
 PN WO200168848-A2.
 XX
 PD 20-SEP-2001.
 XX
 PF 28-FEB-2001; 2001WO-US06520.
 XX
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 03-MAR-2000; 2000US-187202P.
 PR 06-MAR-2000; 2000US-186968P.
 PR 14-MAR-2000; 2000US-189320P.
 PR 14-MAR-2000; 2000US-189328P.
 PR 15-MAR-2000; 2000WO-US06684.
 PR 21-MAR-2000; 2000US-190828P.
 PR 21-MAR-2000; 2000US-191007P.
 PR 21-MAR-2000; 2000US-191048P.
 PR 21-MAR-2000; 2000US-191314P.
 PR 28-MAR-2000; 2000US-192655P.
 PR 29-MAR-2000; 2000US-193032P.
 PR 29-MAR-2000; 2000US-193053P.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 04-APR-2000; 2000US-194449P.
 PR 04-APR-2000; 2000US-194647P.
 PR 11-APR-2000; 2000US-195975P.
 PR 11-APR-2000; 2000US-196000P.
 PR 11-APR-2000; 2000US-196187P.
 PR 11-APR-2000; 2000US-196690P.
 PR 11-APR-2000; 2000US-196820P.
 PR 18-APR-2000; 2000US-198121P.
 PR 18-APR-2000; 2000US-198585P.
 PR 25-APR-2000; 2000US-199397P.
 PR 25-APR-2000; 2000US-199550P.
 PR 25-APR-2000; 2000US-199654P.
 PR 03-MAY-2000; 2000US-201516P.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 05-JUN-2000; 2000US-209832P.
 PR 28-JUL-2000; 2000WO-US20710.
 PR 22-AUG-2000; 2000US-0644848.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 08-NOV-2000; 2000WO-US30952.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000WO-US34956.
 XX
 XX (GETH) GENENTECH INC.
 PA
 PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 XI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 XX
 DR N-PSDB; AAS46223.
 XX
 PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 XX presence of tumours, such as prostate and breast tumours, in mammals and
 XX to screen for modulators of the compounds -
 XX

N-PSDB; AAF92138.

Eighty four nucleic acids encoding PRO polypeptides, useful in molecular biology, including use as hybridization probes, and in chromosome and gene mapping.

Claim 12; Fig 162; 278pp; English.

The present sequence is a human PRO polypeptide (secreted and transmembrane). The PRO protein, and PRO agonists, PRO antagonists or anti-PRO antibodies are useful for preparation of a medicament useful in the treatment of a condition which is responsive to the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO protein may also be employed as molecular weight markers for protein electrophoresis. The PRO coding sequence has applications in molecular biology, including use as hybridisation probes, and in chromosome and gene mapping.

Sequence 705 AA;

Query Match	64.9%;	Score 2339;	DB 22;	Length 705;
Best Local Similarity	66.38;	Pred. No. 4.6e-213;		
Matches 459;	Conservative 66;	Mismatches 133;	Indels 34;	Gaps 7

[illegible]

RESULT 5 . . .
AAB61884 .

ID	AA
AAB61884	standard; Protein; 705 AA

AC : AAB61884 ;

DT	08-MAY-2001	(first entry)
YY		

Chimeric Zcytor14 protein #2

KW Cytokine receptor; zcyto14; human; inflammation; rheumatoid arthritis;
KW antiinflammatory; gene therapy; vaccine.

OS Homo sapiens.

PN WO200104304-A1

PD 18-JAN-2001

PF 30-JUN-2000; 2000WO-US18383

PR 07-JUL-1999; 9905-0348854

PA (ZYMO) ZYMOGENETICS INC.
XX

PI Presnell SR, Burkhead SK, Powder SL,

DR WPI; 2001-112618/12
yy

PT New polypeptide encoding a human cytokine receptor zcytor14, for
PT treating inflammation e.g. rheumatoid arthritis -

PS Claim 2; Page 105-107; 112pp; English

The invention provides a new human cytokine receptor designated zcytor14. zcytor14 can be expressed by standard recombinant methodology. The encoding nucleic acid is useful for detecting the expression of a zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be used to screen biological samples in vitro for the presence of zcytor14. Proteins, polypeptides and peptides having zcytor14 activity can be administered to a subject who lacks an adequate amount of this polypeptide, for treating inflammation and conditions such as rheumatoid arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14 antibodies) can be used to treat a subject who produces an excess of zcytor14. zcytor14 nucleotide sequences can also be used to provide zcytor14 to a subject. The present sequence represents a chimeric zcytor14 protein.

SQ Sequence 705 AA;

Query Match	64.9%;	Score 2339;	DB 22;	Length 705;
Best Local Similarity	66.3%;	Pred. No. 4.6e-213;		
Matches 453; Conservative	66;	Mismatches 133;	Indels 34;	Gaps 7

```

OY 1 MVSNEFLLSIALGRNPVVVSLERLMEPODTRKSGLSCHLMDBGVCLPSTLSAGAPV 60
OY 1 MVSNEFLLSIALGRNPVVVSLERLMEPODTRKSGLSCHLMDBGVCLPSTLSAGAPV 60
Db 1 mpvnpfllsialgrnpvvslsterlrvpdda chqgqslscdlwdsdlloclpdlvpaegpv 60
OY 61 LVPRLOLEVLRCPOKTDICALRVVVVHLAVHGMAREP-----AGKSDELEDSRNAS 116
Db 61 lprhqlqelvlrcqctcdclrlvaahlvahgmwepdeekfggaadsveeaprnas 120
OY 117 LQAGVWLSFOAPRIPARCALLLEVOPADVLOPOGVSAGVAPFCFEASTLGAEOVIMSYPKR 176
OY 117 LQAGVWLSFOAPRIPARCALLLEVOPADVLOPOGVSAGVAPFCFEASTLGAEOVIMSYPKR 176
Db 121 lqagvwlslfgayptarcvlllevqpaaivqfgsgvsyvvdcafealgsevrllwslcyqr 180
OY 177 YKQENLNLQOLP-----DGDNVLLTLQVSEEDQSPFLTLXLRPVDALKSLMYNRLT 227
OY 177 YKQENLNLQOLP-----DGDNVLLTLQVSEEDQSPFLTLXLRPVDALKSLMYNRLT 227
Db 181 ykqenlnhcgqdrpdlpwlinsadagavhylvlnseeghlgslstymnqvgppkprwhnlt 240
OY 228 GQNTYTLNHTDLPVCLCIQWVSLSEDPSEKVEPCFPRDPGAAHRLMHTARLRYVLSPGWQ 28
OY 228 GQNTYTLNHTDLPVCLCIQWVSLSEDPSEKVEPCFPRDPGAAHRLMHTARLRYVLSPGWQ 28

```



```

241 gpgqlltlnhtlvpclclqvwplpdpdsvrtnlcpfcdprahqnlqgaarlrltqlgwsl 300
288 LDAPCCLPFGKWTLCQWADPQSPQPLVPVPVPOKNAVTNPNODEPOLVAGHPNLCVOYSTWE 347
301 ldpocslpaeaalcwtrpagsdcgplvpplswentvtdvletpilkhpnltcvynsse 360
348 KVOLOACSMADSLGPFKDDMLVEMKGTGLNMTSVCLSPSGCTPLPSMASTRARIGEEL 407
361 klqgqecwadsipklkddvlllctcrpqnrlcalpescgclspkskstrairgeyl 420
408 LQDRSHOCMLMNDNNGSLWACPMKXTHRRWVLYWLAACLLAALFFELLKDRRK 467
421 lqdlqsgqclqlw-dddldgalwacpmdkylhkrwalvliacllfaaalsllllkxhak 479
468 -----AARGSRFALLHSADGAGYERLVGALSLSQMLRVAVDLSRRE 513
480 gwllllkqdvsgaaarg-raaillysadsgferlvgalasalcqplrvavdlvsrre 538
514 LSAHGALAMFHRRRLLOEGGVVILFSPAAYAOQOWILOQTVEP---GPHDALAAML 570
539 lsaqgpavafnaqrrqlqeggvvllfsgavalcsewldqgvsqgahghpdafrasl 598
571 SCVLPDFLGKATGKRYGVYFPGDLHPDSVSPFRVAPLFSLEPTQLPAFLDALOGGCS 630
599 scvlpdflqgrapsyvgacldrlhpdavpalfrtvpvftlpsqldflgalqgprapr 658
631 AGRPADRVERVTOALRSALDSC--TSSSEAPG 660
659 sgrlqraeqvsralqpaldsyfnppgtpapg 690

```

RESULT 6

AAB61885 standard; Protein: 675 AA.

AAB61885:

08-MAY-2001 (first entry)

Chimeric zcytor14 protein #3.

Cytokine receptor; zcytor14; human; inflammation; rheumatoid arthritis; antiinflammatory; gene therapy; vaccine.

Homo sapiens.

W0200104304-AL.

18-JAN-2001.

30-JUN-2000; 2000WO-US18383.

07-JUL-1999; 99US-0348854.

(ZYMO) ZYMOGENETICS INC.

Presnell SR, Burkhead SK, Powder SL;

WPI; 2001-112618/12.

New polypeptide encoding a human cytokine receptor zcytor14, for treating inflammation e.g. rheumatoid arthritis -

Claim 2; Page 107-109; 112pp; English.

The invention provides a new human cytokine receptor designated zcytor14. zcytor14 can be expressed by standard recombinant methodology. The encoding nucleic acid is useful for detecting the expression of a zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be used to screen biological samples in vitro for the presence of zcytor14. Proteins, polypeptides and peptides having zcytor14 activity can be administered to a subject who lacks an adequate amount of this polypeptide, for treating inflammation and conditions such as rheumatoid

CC arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14 CC antibodies) can be used to treat a subject who produces an excess of zcytor14. zcytor14 nucleotide sequences can also be used to provide CC zcytor14 to a subject. The present sequence represents a chimeric CC zcytor14 protein.

SQ Sequence 675 AA;

Query Match 63.2%; Score 2277; DB 22; Length 675;
 Best Local Similarity 65.9%; Pred. No. 3,4e-207;
 Matches 447; Conservative 63; Mismatches 132; Indels 36; Gaps 6;

```

QY 1 MPVSFLLSLGRNPVVYSLEIRMEPODTRFARCSIGLSCHIMQDVLCIPGSLQSPAPV 60
DB 1 mpvsvfllslagrnpvvysleirmevpoetrarfarsiglschmqdvllcpgdvlvppgv 60
QY 61 LVPTFLQTELVLRCPQKDCALRVRRVVLAVGHAAEE---AGKSDSELOESRNAS 116
DB 61 lvptflqtelvllrcpqkdcclrvavvlavghaeedekfsgaadsyveepnas 120
QY 117 LQAOVLSFOAYPIARCALLEVQVPADLVQPGQSGAVFDCFEASLGAEOIWSYTKPR 176
DB 121 lqaovlsfogaipiarcallevqvpadlvqpgqsgavvydcfeaalgservlwsytqpr 180
QY 177 YQKELNLTQOLP-----DGDNVLLFDYSEEDDFELLYLRVPDALKSLMKNL 227
DB 181 yqelnlntqolp-----dgdnvllfdyseeddfellylrvpdalkslmknlt 240
QY 228 GPONTTLNHTLVPCICIQWVSLSPDSEVEREFCFEREDGAAHRLHRIARLSPGWQ 287
DB 241 gponltlnhtlvpclclqvwplpdpdsvrtnlcpfcdprahqnlqgaarlrltqlgwsl 300
QY 288 LDAPCCLPFGKWTLCQWADPQSPQPLVPVPVPOKNAVTNPNODEPOLVAGHPNLCVOYSTWE 347
DB 301 ldpocslpaeaalcwtrpagsdcgplvpplswentvtdvletpilkhpnltcvynsse 360
QY 348 KVOLOACSMADSLGPFKDDMLVEMKGTGLNMTSVCLSPSGCTPLPSMASTRARIGEEL 407
DB 344 klqgqecwadsipklkddvlllctcrpqnrlcalpescgclspkskstrairgeyl 403
QY 408 LQDRSHOCMLMNDNNGSLWACPMKXTHRRWVLYWLAACLLAALFFELLKDRRK 467
DB 404 lqdlqsgqclqlw-dddldgalwacpmdkylhkrwalvliacllfaaalsllllkxhak 462
QY 468 AARGSRFALLHSADGAGYERLVGALSLSQMLRVAVDLSRRELASHAGALAMFHRR 527
DB 463 aargsrfaallhsadsgferlvgalasalcqplrvavdlvsrrelsaqgpavafnaqr 522
QY 528 RRILOEGGVVILFSPAAYAOQOWILOQTVEP---GPHDALAAMLSCVLPDLOGRAT 584
DB 523 rrlloeggvvllfsgavalcsewldqgvsqgahghpdafraslscvlpdflqgrap 582
QY 585 RYGVYFPGDLHPDSVSPFRVAPLFSLEPTQLPAFLDALOGGCSGSRPADRVERVTOA 644
DB 583 ryvgvfpdglhpdsvspfrvaplfsleptqlpafladaloggcsstgpadrvervtoa 642
QY 645 LRSAIDSC--TSSSEAPG 660
DB 643 lrpaldsyfnppgtpapg 660

```

RESULT 7

AAB61883 standard; Protein: 688 AA.

AAB61883:

08-MAY-2001 (first entry)

Chimeric zcytor14 protein #1.

Cytokine receptor; zcytor14; human; inflammation; rheumatoid arthritis;

antiinflammatory; gene therapy; vaccine.

Homo sapiens.

MO200104304-A1.

18-JAN-2001.

30-JUN-2000: 2000MO-US18383.

07-JUL-1999: 99US-0348854.

(ZYMO) ZYMOGENETICS INC.

Presnell SR, Burkhead SK, Powder SL.

WPI: 2001-112618/12.

New polypeptide encoding a human cytokine receptor designated zcytor14, for treating inflammation e.g. rheumatoid arthritis -

Claim 2: Page 102-104; 112pp: English.

The invention provides a new human cytokine receptor designated zcytor14. zcytor14 can be expressed by standard recombinant methodology. The encoding nucleic acid is useful for detecting the expression of a zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be used to screen biological samples in vitro for the presence of zcytor14. proteins, polypeptides and peptides having zcytor14 activity can be administered to a subject who lacks an adequate amount of this polypeptide, for treating inflammation and conditions such as rheumatoid arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14 antibodies) can be used to treat a subject who produces an excess of zcytor14. zcytor14 nucleotide sequences can also be used to provide zcytor14 to a subject. The present sequence represents a chimeric zcytor14 protein.

Sequence 688 AA:

Query Match 62.8%; Score 2263.5; DB 22; Length 688;
Best Local Similarity 64.9%; Pred. No. 6,7e-206;
Matches 449; Conservative 63; Mismatches 129; Indels 51; Gaps 8;

1 MVSFSLSLALGKRNPNVSLERLMEPODTRARCSLGLSHMDGVLCIPGLQSAAGCV 60
1 mvpwflslslalgrspvslslrlvpgdchpcpsjrlwdsdlclpgdlvpgpvp 60
61 LVPTRLTETLVRCPOKTDALRYRVVHVLAVHGMARPEE---AGKSDSELOESRNAS 116
61 laptlltqetvlvtctqkcdclclrvavhavlhwpeedeekfgsdaadsveeprnas 120
117 LQAQVYLSFQAPYPIARCALLEVQPADLVQPGSVSAVFDFCEASLGAELVIMWSTKPR 176
117 lqgqvylsfqapyparcvlllevqpaalvtqfgsvgyvdfcealgselvwstlqpr 180
177 YQKELNLTQOLP-----DGNVLTLDVSEQDPSFLIYLPRVPALKSLWTKNLT 227
177 yqkelnltqolp-----dgnvltldvseodpsfllylprvpalkslwtknlt 227
181 yekelnltqolpavlvsadgnvhlvlnvseeqhfslsywngvqgpprvphknt 240
228 GPQNTLNHNDVLCICIQVWSLEPDSRVFECPRPDGARNHMTARLAVLSGVQV 287
228 gpqntlnhndvclciqvwslepdsrvfecprpdgarhmtarlavlsgvqv 287
241 gpgqtltnhndvclciqvwslepdsrvfecprpdgarhmtarlavlsgvqv 300
288 LDAPFCLPGKVTLCMOAPDSCOPLYPVPPOKNAVNEPODFOLVAGHPNLGVQVSTWE 347
288 ldapfclpgkvtlcmoapdscoplypvpvpoKNVNEPODFOLVAGHPNLGVQVSTWE 347
301 ldpapclpaaealewrepqgpcplvppiswenvtd-----vnse 343
348 KVOQLQACSMADSLGPFKDDMLLVEMKTLNNTSVCALEPDSGCTPLPSMASTRARLGEEL 407
348 kvqlqacsmadslgpfkddmllvemktlntsvcalepdsGCTPLPSMASTRARLGEEL 407
344 klqgqelwadsipklkdvlllletrgpdnrlcalpsgtslpskstratrlgeyl 403
408 LQDFRSQCQMLMDNDNGSLMACPMKXYIHRRWLVWLAACLLLAALFFLLKDKDRR 467

DB 404 lqdlsgsgcqlw-dddlgalwacpmckylhkrwalvwlacillfaaalslilllkkdnak 462
QY 468 -----AANGSTRALLHSADGAGYERLVGALASALSQMPLRVAVDMSRE 513
DB 463 gwrlrlkqdvrsaaag-raalllvsadsgferlygalasalcqplrvavdlwstte 521
QY 514 LSAHGALAMFHHORRRILOEGVILLFSPAANAOCQOMLOQTEP---GPHDALAAML 570
DB 522 lsaqgpvaawfhagrtqrlqggvvvllfsgaavalcsewldqvsqgphabphdfrsl 581
QY 571 SCVLPDFLAGRATGRVGVYFDFGLHDPDSVSPRRVADFLPQLPAPFLDALGCGCSTS 630
DB 582 scvlpdfllgrrapsyvgacfdrlhpdavpalftvptflpsqldpflgalqgprapr 641
QY 631 AGRADRVERTQALRSALDSC---TSSEAPG 660
DB 642 sgrlqeraeqvrsalqpaldsyfhnpptlpap 673

RESULT 8

ID AAB61881 standard; Protein: 575 AA.

AC AAB61881;

DT 08-MAY-2001 (first entry)

DE Human variant zcytor14 protein zcytor14-1.

KM Cytokine receptor; zcytor14; human; inflammation; rheumatoid arthritis;
antiinflammatory; gene therapy; vaccine; variant; zcytor14-1.

OS Homo sapiens.

PN MO200104304-A1.

PD 18-JAN-2001.

PF 30-JUN-2000: 2000MO-US18383.

PR 07-JUL-1999: 99US-0348854.

PA (ZYMO) ZYMOGENETICS INC.

PT Presnell SR, Burkhead SK, Powder SL;

PT WPI: 2001-112618/12.

PT N-PSDB: AAC85029.

PT New polypeptide encoding a human cytokine receptor zcytor14, for
treating inflammation e.g. rheumatoid arthritis -

PS Disclosure: Page 2-3; 112pp: English.

XX The invention provides a new human cytokine receptor designated zcytor14.
zcytor14 can be expressed by standard recombinant methodology. The
encoding nucleic acid is useful for detecting the expression of a
zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be
used to screen biological samples in vitro for the presence of zcytor14.
proteins, polypeptides and peptides having zcytor14 activity can be
administered to a subject who lacks an adequate amount of this
polypeptide, for treating inflammation and conditions such as rheumatoid
arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14
antibodies) can be used to treat a subject who produces an excess of
zcytor14. zcytor14 nucleotide sequences can also be used to provide
zcytor14 to a subject. The present sequence represents a variant of
the human cytokine receptor zcytor14, designated zcytor14-1. This
variant is a truncated form of the receptor polypeptide and lacks
amino acid residues 1-113 of zcytor14.

Sequence 575 AA:

PR 12-MAY-1998; 98US-0085094.
PR 12-MAY-1998; 98US-0085105.
PR 12-MAY-1998; 98US-0085180.
PR 18-MAY-1998; 98US-0085906.
PR 18-MAY-1998; 98US-0085920.

PR 18-MAY-1998; 9805-0085921.
PR 18-MAY-1998; 9805-0085922.

PR	18-MAY-1998;	98US-0085924.
PR	18-MAY-1998;	98US-0085924.
PR	18-MAY-1998;	98US-0085924.

PR	18-MAY-1998;	98US-0085925.
EN	20 MAY 21/50;	0000 00000000
10	10 MAY 10/00	0000 00000007

XX
DA

HUMAN GENOME SCI INC
(HUMA-)

XX
PI Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;

PI	Olsen HS,	Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;
PI	Lafleur DW,	Endress GA, Edner R;

XX
DR WPI; 2000-062296/05.

XX	N-FSDB, A5A0JZ05.
UR	A5A0JZ05.

useful for diagnosis and treatment of e.g. cancers, neurological diseases, immune disorders, inflammation or blood disorders -

XX
claim 11. page 372-373. 475no: English.
DS

XX
AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes

CC AAY76124 to AAY76223 are the secreted proteins encoded by the 9' numal
CC genes. The gene encoding this protein was found to be on chromosome 3

CC The genes and their corresponding secreted polypeptides are
CC useful for preventing, treating or ameliorating medical conditions,

cc diagnosed by determining the amount of the new polypeptides in a sample.

uses are described for each of the 97 genes, based on which tissues the genes are most highly expressed in, and include developing products for the

CC diagnosis or treatment of cancer, tumours, developmental abnormalities,
CC and foetal deficiencies, blood disorders, diseases of the immune system

disorders, autoimmune diseases, schizophrenia, inflammatory diseases, allergies, Alzheimer's and cognitive disorders, arthritis, asthma, psoriasis, sepsis, skin

disorders, digestive/endocrine disorders, infections and AIDS. The

The sequences shown in AY76224 to AY76424 represent fragments of the

700

Query Match	Score 958.5; DB 21; Length 309;
Best Local Similarity	64.68; Pred. No. 2.4e-82;

Matches	192;	Conservative	20;	Mismatches	50;	Inners	29;	Gaps
---------	------	--------------	-----	------------	-----	--------	-----	------

0Y
1 MFVSWLHDSLMONKVFVVSJUNWMLFQZANMVCBOGCHNDODVMOZUOZGOMV
||| ||| ||| : ||| : ||| | ||| : |||
||| ||| ||| : ||| : ||| | ||| : |||

[illegible][illegible]

117 1CAQVNI-SFOAYPIARCALLEVOVPADLVOPGOSVCSAVFDCFEASLGAEOQIMSYTKR 176

Db 121 lgaqvlsfdayptarcvllvqvpaalvqfgsgvsgvvydcfeaalgsevrimsytqpr 180

QY 177 YOKELNTQOLP-----DCNNLLTDVSEEDFSFLYL 212

```

181 yekexnhtqglpdcrglewinsipscwalpwlInvSadgdnvhlvlnvseeghfjlslwv 240
db

```

QY 213 PVPDAKSLWYKNLTGPONITLHNTDLVPCLCIQVMSLEBDESEVEFCPFREDDPGAH 269

261 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

RESULT 10

AM25864 standard; Protein; 332 AA.

AM25864;

16-OCT-2001 (first entry)

Human protein sequence SEQ ID NO:1379.

Human; cancer; ulcer; HIV infection; human immunodeficiency virus; anti-inflammatory; antineoplastic; antirheumatic; immunosuppressive; antibacterial; endocrine; cardiac; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia; dermatologic; haemostatic; vulnery; antilucer; osteopathic; eczema; neuroprotective; antidepressant; nootropic; antiparkinsonian; infection; immunostimulant; gene therapy; antisense therapy; vaccine; inflammation; antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis; cardiac dysfunction; neuropathologic; cardiac anaphylaxis; autoimmunity; genetic disease; hematopoietic disorder; platelet disorder; asthma; thrombocytopaenia; osteoporosis; severe combined immunodeficiency; allergic rhinitis; diabetes; multiple sclerosis; depression; Alzheimer's disease; Parkinson's disease; neurodegenerative disorder; neurological disorder.

Homo sapiens.

MO200153455-A2.

26-JUL-2001.

22-DEC-2000; 2000MO-0535017.

23-DEC-1999; 99US-0471275.

21-JAN-2000; 2000US-0488725.

25-APR-2000; 2000US-0552317.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI: 2001-457603/49.

N-PSDB; AAH99805.

Isolated human polynucleotides encoding polypeptides, useful for the treatment and diagnosis of e.g. cancer, ulcers and HIV infection.

Claim 20; Page 283; 1217pp; English.

AAH99166 to AAH99904 encode the human proteins given in AAH25225 to AAH25963. The proteins can have activities based on the tissues and cells they are expressed in, such as: anti-inflammatory; antirheumatic; antiallergic; immunosuppressive; antibacterial; endocrine; cardiac; cardiovascular; antianaemic; antilucer; osteopathic; dermatologic; antiallergic; antianaphylactic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for disorders or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, hematopoietic disorders, anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and

CC neurological disorders.

XX Sequence 332 AA;

Query Match 23.3%; Score 841.5; DB 22; Length 332;
Best Local Similarity 62.1%; Pred. No. 3.7e-71;
Matches 169; Conservative 22; Mismatches 52; Indels 29; Gaps 3;

1 MPVSFELSLALGRNPVYVSLERLMEODPARCSLGLSCHLMDVDVCLGSLQASAGPV 60

2 MPVWFLSLALGRNPVYVSLERLMEODPARCSLGLSCHLMDVDVCLGSLQASAGPV 61

61 LVPTRLQELVLRPCQKTCALRVRVHLAVGHNAEPE-----AGKSESELOESNRAS 116

62 LAPTQLQELVLRPCQKTCALRVRVHLAVGHNAEPE-----AGKSESELOESNRAS 121

117 LQAVYLSQAVPIARCALLERQVDPDVGQSGVSAVFDCFEASIGAEVQIMSTYKPR 176

122 LQAVYLSQAVPIARCALLERQVDPDVGQSGVSAVFDCFEASIGAEVQIMSTYKPR 181

177 YQKELNLRQQLP-----DGDVYLLTLDVSEEDSFLLKLR 212

182 YQKELNLRQQLP-----DGDVYLLTLDVSEEDSFLLKLR 244

213 PVPDALKSLMYNLTGPONTITNHTDVLVPCLC 244

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

242 qvqppkprwhknlcpqlltlnhtdlypolcigwplpdsrvtsapsgrfp-ah 296

Query Match 8.6%; Score 310.5; DB 22; Length 617;
 Best Local Similarity 25.7%; Pred. No. 3.1e-20;
 Matches 123; Conservative 55; Mismatches 208; Indels 93; Gaps 17;

224 KNLGPGQITLNTLTPCLCIGWMSLEPDSERVERCPPEPCAH-RNLWHIARLVLS 282
 183 klvsghvtdlypeflipcmcleasylgedltvrrkcfqspwpaesdflvgsitfdys 242
 283 PG---VMQLDAPCCPGKVTLCMQADPSPCOPLVPPVPPKKNATVNEPQDQVAG---H 336
 243 qhngmvmaltlrcplkleaslcwrgqlpcpceli-----pnatagesegwyllenvdlh 296
 337 PNLGVQSTWEKVVQAC-----SMADSLGPFKDDMLLVEMKGTGLNNTSVCALEPSG 388
 297 pqlcfkfs-fensnvechpgsgslpswtvsmtd-qaqtlhtsrtlyafsaawdpg 354
 389 C---TPLESMASSTRARLGEELLQD-----FRSHQCMQWMDNMGSLMACPMDKYIHR 440
 355 lpgdtmppyyslsgtqgsvpytldllpflrgencillvrsd-----vifa 401
 441 WVLY-----WLAQLLAAALFFLLKKDRKARAGSTRALLHSADAGYERLVAL 493
 402 wkhvlpddapytqlil-----tslgsgrtlpvlllhaadseaqrllvga1 448
 494 ASALSOM---PLRVAVDLSRRELNAHGALEWFMHQRRIIEGCVILLESPPAVAQCC 550
 449 aellrtalgggrdvldwqeghvarlgrlpvwaarervaregvtvlllmcagpps--- 505
 551 QWLOLQVPEPQHDLAAMSCVLPDELOGRATGRVGVYFDGLHPDSVSPRVALEF 610
 506 ----tacsqgpgaastrlllcaaprpl-----llayfsrlcaxgdldprlralpry 552
 611 SLPTQLPAFLALAGGCSNSA-----GRPADRVERVTA-----LRNALDSCTSSSEAP 659
 553 rllrdlprllraldaqpaltlaaswshlgakrclknlegchilleleaakdyggsstsp 611

RESULT 15

AAU04957 standard; Protein; 667 AA.

AAU04957;

24-OCT-2001 (first entry)

Human Interleukin 17 receptor, IL-17RH3.

Human: Interleukin-17 receptor; IL-17RH3; agonist; antagonist;
 PRO9877; DNA 119502-2789; systemic lupus erythematosus;
 rheumatoid arthritis; osteoarthritis; diabetes mellitus;
 allergic disease; asthma; demyelinating disease;
 degenerative cartilaginous disorder; transplantation associated disease.

Homo sapiens.

Key Location/Qualifiers
 Peptide 1..23
 Protein /label- Signal_peptide
 24..667
 Region /label- Mature_IL_17RH3
 90..96
 Region /note- "N-myristoylation site"
 104..108
 Region /note- "cAMP/GMP-dependent protein kinase
 phosphorylation site"
 318..322
 Modified-site /note- "Asn is N-glycosylated"
 322..329
 Region /note- "Tyrosine kinase phosphorylation site"
 347..351
 Modified-site /note- "Asn is N-glycosylated"
 358..364
 Region

FT /note- "N-myristoylation site"
 FT 364..368
 FT /note- "Asn is N-glycosylated"
 FT 453..462
 FT /note- "Eukaryotic cobalamin-binding protein motif"
 FT 455..472
 FT /note- "Transmembrane domain"
 FT 470..476
 FT /note- "N-myristoylation site"
 FT 482..486
 FT /note- "Glycosaminoglycan attachment site"
 FT 645..649
 FT /note- "cAMP/GMP-dependent protein kinase
 phosphorylation site"

MO200146420-A2.

28-JUN-2001.

20-DEC-2000; 2000MO-US34956.

23-DEC-1999; 99US-0172096.

30-DEC-1999; 99MO-US31274.

11-JAN-2000; 2000US-0175481.

18-FEB-2000; 2000MO-US04341.

02-MAR-2000; 2000MO-US05841.

21-MAR-2000; 2000US-0191007.

21-MAR-2000; 2000MO-US07532.

02-JUN-2000; 2000MO-US15264.

22-JUN-2000; 2000US-0213087.

22-AUG-2000; 2000US-0644848.

24-AUG-2000; 2000MO-US23338.

24-OCT-2000; 2000US-0242837.

10-NOV-2000; 2000MO-US30873.

28-NOV-2000; 2000US-0253646.

01-DEC-2000; 2000MO-US32678.

(GETH) GENENTECH INC.

Chen J, Filvaroff E, Fong S, Goddard A, Godowski PJ, Grimaldi CJ,

Gurney AL, Li H, Hillan KJ, Tumas D, Van Lookeren M, Vandlen RL,

Watanabe CK, Williams PM, Wood WI, Yansura DG;

WPI; 2001-451708/48.

N-PSDB; AAS09516.

Novel PRO polypeptides homologous to interleukin-17, useful for the

diagnosis and treatment of immune related disease e.g. rheumatoid

arthritis and diabetes

Claim 10; Fig 16; 188pp; English.

The sequence is PRO9877 which is the human interleukin 17 receptor,
 IL-17RH3, encoded by DNA 119502-2789. A composition
 containing ant/agonists to the PRO polypeptides or individual components
 are useful for treating a mammal with an immune related disease e.g.
 systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
 juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
 idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
 vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
 thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
 disease, a demyelinating disease, an autoimmune or immune-mediated skin
 disease, contact dermatitis, an allergic disease e.g. food
 hypersensitivity, asthma, a transplantation associated disease, or a
 chronic inflammatory demyelinating polyneuropathy. Treating a
 degenerative cartilaginous disorder comprises administering a PRO1031 or
 PRO1122 polypeptide agonist, or antagonist to the mammal. Numerous
 examples of the diseases and disorders are given in the specification.

Sequence 667 AA;

Query Match

8.6%; Score 310; DB 22; Length 667;

Best Local Similarity 28.1%; Pred. No. 3,9e-20;
Matches 130; Conservative 56; Mismatches 214; Indels 62; Gaps 18;

```
224 KNLTPONTLTHNTDVLPCICIQWLSLEPDSERVECPREDPGAH-RNLW--HIALRL 279
223 KIVSGHVELPVEFLIPCLCEASYLQEDTVRRKKCFGSPPEAYGDFWKSHTDYS 282
280 VLSFGWOLDAFCCLPGKVTLCWQAPDQSPCOPLVPPVPOKNATVNEPQDF--QLVAGH 336
283 qhtgmymaltlrcplklealecgrhndhclckdl-----pnataresdgwyvylekvdllh 336
337 PNLGVOSTWKEVQLOA-----CSWADSLGPEKDDMLVEMKGLNNTSYCALEPSCG 389
337 pqlckkfsfgnshvecphtqtsltswnvsmtdt-qaqqlilhfsrmhatfsaawslpgl 395
390 ---PFLPSMASTRRAARLGEELLAD-----FRSHOCQQLWMDNDNMGSLW---ACPMCKYTH 438
396 gqdtlvppvytvsgarqsspsvslidllipflrpgccvltwrsd-vqfawkhlcpdvsyrh 454
439 RRVVLWVLACLLIAALFFFLILKDRRKARG-SPTALLHSADGAGYERLYGALASNL 497
455 Ig--lllallallcllgyvalctrrpgsgpparpvlllhaadseagrtlvgaiaell 512
498 SOM--PLRAVADLMSRRELSAHGALAMFHRRRIQEGGVILLFSPAAVAQCOQWLQ 554
513 raalggrdrvlvdlwegrlnvarvqplwlaartlvareqtvlllwsga-----d 563
555 LQYVEGPHDALAANLSCVLPDFLOGRATGRYGVGVPDGLHPDSVSPFRVAPLESLEPT 614
564 lrps-gpdptraaplla-----llhaaprpilllayfsrlcakgdippralpryllr 617
615 QLPALFDALOG---GCSTAGRPADREVERTQALRSALDSCT 653
618 dlprllraldarfaeatwgrlgarqr-----qsrlelcs 654
```

Search completed: September 28, 2002, 19:39:44
Job time: 6372 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

CM protein - protein search, using sw model

Run on: September 28, 2002, 17:55:57 ; Search time 27.72 Seconds

(Without alignments)
593.898 Million cell updates/sec

Title: US-09-899-471-2

Perfect score: 3605
1 MPVSWFLSLALGRNPVVS.....SSEAPCCCEMDGPPCTTLE 674

Sequence:

BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2.6/ptodata/2/1aa/5a_COMB.pep.*
- 2: /cgn2.6/ptodata/2/1aa/5a_COMB.pep.*
- 3: /cgn2.6/ptodata/2/1aa/5a_COMB.pep.*
- 4: /cgn2.6/ptodata/2/1aa/5a_COMB.pep.*
- 5: /cgn2.6/ptodata/2/1aa/5a_COMB.pep.*
- 6: /cgn2.6/ptodata/2/1aa/5a_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	310.5	8.6	617	4	US-09-188-930-303
2	167	4.6	864	2	US-08-620-694A-2
3	167	4.6	864	3	US-09-022-255-2
4	167	4.6	864	3	US-09-022-255-2
5	167	4.6	864	3	US-09-022-255-2
6	167	4.6	864	3	US-09-022-255-2
7	167	4.6	864	3	US-09-022-255-2
8	167	4.6	864	3	US-09-022-255-2
9	167	4.6	864	3	US-09-022-255-2
10	151.5	4.2	866	2	US-08-620-694A-10
11	151.5	4.2	866	3	US-09-022-255-10
12	151.5	4.2	866	3	US-09-022-255-10
13	151.5	4.2	866	3	US-08-978-773-4
14	151.5	4.2	866	3	US-09-022-253-10
15	151.5	4.2	866	3	US-09-022-253-10
16	151.5	4.2	866	3	US-09-022-253-10
17	151.5	4.2	866	3	US-09-022-253-10
18	151.5	4.2	866	3	US-09-022-253-10
19	149	4.1	330	4	US-09-188-930-125
20	102.5	2.8	1498	2	US-08-404-531B-28
21	102.5	2.8	1498	3	US-08-404-531B-28
22	102.5	2.8	1498	3	US-08-404-531B-28
23	102.5	2.8	1498	3	US-08-404-531B-28
24	102.5	2.8	1498	3	US-08-404-531B-28
25	102.5	2.8	1498	3	US-08-404-531B-28
26	100	2.8	1581	3	US-08-476-900A-6
27	100	2.8	1581	3	US-08-476-900A-6
28	100	2.8	1581	3	US-08-476-900A-6
29	100	2.8	1581	3	US-08-476-900A-6
30	100	2.8	1581	3	US-08-476-900A-6
31	100	2.8	1581	3	US-08-476-900A-6
32	100	2.8	1581	3	US-08-476-900A-6
33	100	2.8	1581	3	US-08-476-900A-6
34	100	2.8	1581	3	US-08-476-900A-6
35	100	2.8	1581	3	US-08-476-900A-6
36	100	2.8	1581	3	US-08-476-900A-6
37	100	2.8	1581	3	US-08-476-900A-6
38	100	2.8	1581	3	US-08-476-900A-6
39	100	2.8	1581	3	US-08-476-900A-6
40	100	2.8	1581	3	US-08-476-900A-6
41	100	2.8	1581	3	US-08-476-900A-6
42	100	2.8	1581	3	US-08-476-900A-6
43	100	2.8	1581	3	US-08-476-900A-6
44	100	2.8	1581	3	US-08-476-900A-6
45	100	2.8	1581	3	US-08-476-900A-6

28	100	2.8	1410	4	US-09-567-969-3	Sequence 3, Appl1
29	100	2.8	1410	4	US-09-568-480-3	Sequence 3, Appl1
30	100	2.8	1410	4	US-09-568-480-3	Sequence 3, Appl1
31	100	2.8	1410	4	US-09-568-480-3	Sequence 3, Appl1
32	100	2.8	1410	4	US-09-568-480-3	Sequence 3, Appl1
33	100	2.8	1410	4	US-09-568-480-3	Sequence 3, Appl1
34	100	2.8	1410	4	US-09-568-480-3	Sequence 3, Appl1
35	99.5	2.8	1410	4	US-09-568-480-3	Sequence 3, Appl1
36	99.5	2.8	1410	4	US-09-568-480-3	Sequence 3, Appl1
37	97.5	2.7	1618	1	US-07-853-913-4	Sequence 4, Appl1
38	97.5	2.7	1618	1	US-07-853-913-4	Sequence 4, Appl1
39	97.5	2.7	1618	1	US-07-853-913-4	Sequence 4, Appl1
40	95.5	2.6	1618	1	US-07-853-913-4	Sequence 4, Appl1
41	94.5	2.6	1618	1	US-07-853-913-4	Sequence 4, Appl1
42	94	2.6	1618	1	US-07-853-913-4	Sequence 4, Appl1
43	92.5	2.6	1618	1	US-07-853-913-4	Sequence 4, Appl1
44	92.5	2.6	1618	1	US-07-853-913-4	Sequence 4, Appl1
45	91.5	2.5	1618	1	US-07-853-913-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-188-930-303
Sequence 303, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Schachan, Lorna
APPLICANT: Sileman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011C1
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 303
LENGTH: 617
TYPE: PRT
ORGANISM: Mouse
US-09-188-930-303

Query Match 8.6%; Score 310.5; DB 4; Length 617;
Best Local Similarity 25.7%; Pred. No. 1.5e-24;
Matches 123; Conservative 55; Mismatches 208; Indels 93; Gaps 17;

224 KNLGPGQITLNTHTDIPVLCIOVMSLEPDSERFCEPREDPEAH-RNLNHTIARLVIS 282
183 KIVSGGHTVDLEPELPCMCIEASYLOEDTVRKKGKPSWPEAYGSDFOQSIPTDYS 242
283 PG---VMDLAPCCJPGKVTLCMOAPDSCOPVPPQKNATVNEPQDQIVAG---H 336
243 QHNOVVALTLKPLKLESLCKRQDPLTFCET---FNATAOESGVTILENVDLH 296
337 PNLGVYSTWKEVQLOAC-----SMADSLAPFKQDMLEVENKGTINTSVGALBESG 388
297 POLCFKES-FENSSHVECHQSGSLPSMTVSMOT-QAQQTLTHRSSRYATFASAMSDPG 354
389 C---TPLESMSTRARAGEELLQD-----FRSHOQMLNDMDMGSMACPMQYTHRR 440
355 LGPTMPVPIVSIQSGVPTLDTLIDPLRQENCLTVRSR-----VHFA 401
441 WVLV-----WACILALAALEFFLLKKRRKARASFTALLSHSAGAGYELVGL 493
402 WKHVLCPDPAVYPTQLL-----RSLGSGRTIPVLLIHAADSQAQRVLVGL 448
494 ASALSQW---PLRAVVDLMSRRELGAHGLAMPFHORRRILQEGVYILFSPAAYACQ 550

DB 449 AELLRTALGGGRVIVDLMESTHVARIGPLPWLMAAREVAREOGIVLLIMNCAGPS--- 505
 QY 551 QMLQLOTEBEPHDAIAMLSCVLPDLQGRATGRVGVYFDGLLHPDVSPPFRVAPLF 610
 DB 506 -----TACSGDDPOAASLRTILCAAPRPL-----LTAFFSLCAKGIIPRLRALPRY 552
 QY 611 SLPTQAPALDAQGGGSGTA-----GRPADRVRYTQA-----LRSALDSCSTSSSEAP 659
 DB 553 RLRLDPLRLRALDAOPATLASSMSHLGAKRCLKNRLECHLLELAADDDYGGSTNSP 611

RESULT 2

US-08-620-694A-2
 Sequence 2, Application US/08620694A
 Patent No. 5869286

GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin
 APPLICANT: Spriggs, Melanie
 APPLICANT: Fanslow, William
 TITLE OF INVENTION: No. 5869286e1 Receptor That Binds IL-17
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Immunex Corporation
 STREET: 51 University Street
 CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: Apple Operating System 7.5.5
 SOFTWARE: Microsoft Word for Apple, Version 6.0.1
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/620,694A
 FILING DATE: 21 MARCH 1996
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/538,765
 FILING DATE: 7 AUGUST 1995
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/410,535
 FILING DATE: 23 MARCH 1995
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne
 REGISTRATION NUMBER: 34,695

REFERENCE/DOCKET NUMBER: 2617-B
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)587-0430
 TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:

LENGTH: 864 amino acids
 TYPE: amino acid

TOPOLOGY: linear
 MOLECULE TYPE: protein

US-08-620-694A-2

Query Match 4.6%; Score 167; DB 2; Length 864;
 Best Local Similarity 20.3%; Pred. No. 1.1e-08;
 Matches 105; Conservative 62; Mismatches 167; Indels 104; Gaps 25;

DB 224 KNLUT--GPNITLNL-----HNDLVPCLCIQWSLEPD-----SSR 256
 DB 66 KNLUTSSPNITNITVSSSTOHGELVPLVHVE-WTLOTDAIILYEGAEIAYLQINTNER 124
 DB 257 --VEFCPEFREDGAGHNLNHIARLR-VLSPG-----VMOADAPCCLEGGKYLCKQAPQ 307
 DB 125 LCVK--QELSMLOHKKRRKRRFSHFVYVDGQETVTVHHLKPK-----IPDG 171

QY 308 SPCOP-----LVPVPOKNAATVNEPDEFQVLVAGHPNLCYOVSTWKEVLOACSWADSLGPP 363
 DB 172 DPNHRSKIIFVPOEDSKMKMTSCVSSGLMDPNITVETLDTQHLRVDFITLMNEST-PY 230
 QY 364 KDDMLVEAKTGLNTSYV-----CA-----LEP-- 386
 DB 231 Q---VLLSEFSDESNHSCFDVVKOIFAPROEEFHQRANVTFLSKFHHCHHHVQVQPF 287
 QY 387 SGC-----TPLPMSASTRAARLGEELLQDRSRHOCQMLMDDNNGSLMACMDKY 436
 DB 288 SSCINDCLNHAATVPCPVISTNTYVK-----PVADY 318
 QY 437 IHRRWV--LWVLAALLAALFFPL-----LLKDRKRAAGS----- 472
 DB 319 I-PLWYGLITLAILLVGSYIVLILCMFTWLSGADQEKHGDCKINGILLPVADLTPLPL 377
 QY 473 --RTALLHSADGAGTERLYGALASAL-SQMPLRVAVDLSMREELSANGALAMFHQRR 529
 DB 378 RPRKWIIVYSADHPLVYEVVLKFAQFLITACGTEVALDLLEQVISEVGMVMSKOE 437
 QY 530 ILQEGGVILFS-----PAVAQCQOQLQLOTEBEPHDAIAMLSCVL 574
 DB 438 WYENSKITILICSRGTOAKKAILGMAPPAVOLRCDDH-----KPA-GDLFTAAAMNML 490
 QY 575 PDLQGRATGRVGVYFDGLLHPDVSPPFRVAPLFSL 612
 DB 491 PDFKRPACFGTYVVCYFSGICSEHDVDFLNTISRYPL 528

RESULT 3

US-09-022-255-2
 Sequence 2, Application US/09022255
 Patent No. 6072033

GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin
 APPLICANT: Spriggs, Melanie
 APPLICANT: Fanslow, William
 TITLE OF INVENTION: No. 6072033e1 Receptor That Binds IL-17
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Immunex Corporation
 STREET: 51 University Street
 CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZIP: 98101

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: Apple Power Macintosh
 SOFTWARE: Microsoft Word for Apple, Version 6.0.1
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/022,255
 FILING DATE:
 CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/620,694
 FILING DATE: 21 MARCH 1996
 APPLICATION NUMBER: USSN 08/538,765
 FILING DATE: 7 AUGUST 1995

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USSN 08/410,535
 FILING DATE: 23 MARCH 1995

ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia Anne
 REGISTRATION NUMBER: 34,695

REFERENCE/DOCKET NUMBER: 2617-B
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)587-0430
 TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:

LENGTH: 864 amino acids

TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 JS-09-022-255-2

Query Match 4.6%; Score 167; DB 3; Length 864;
 Best local similarity 20.3%; Pred. No. 1,1e-08;
 Matches 105; Conservative 62; Mismatches 167; Indels 184; Gaps 25;

224 KNLN-GRONTLN-----HTDLPCLCIGVSLSPD-----SER 256
 66 KNLPPSPKNTIYINLSVSTGELVPLVHVE-WTLOTASILYLEGAEISVLQINTNER 124
 257 --VECFPRREDPGAHNRNHLRLR-VLSPG-----VMQDAPCCLPKVTLCWQAPDQ 307
 125 LCVKF-QFLSMLOHRRKRFRFSHFVDPGQGEYEVTHLPRK-----IPDG 171
 308 SPQCP-----LVPPVQKNATVNEPQDFQVLVAGHPNLCVOYSTWEKVLQACSWADSLCP 363
 172 DPNKSKITIEVPDCEDESKMKTTSCVSSGLMDPNITVETLDTQHLKVDFTLMNEST-PY 230
 364 KDMLVEMKGTGLNNTSY-----CA-----LEP-- 386
 231 Q---VLLESFSDSEHSCFEDVVKQIFAPRQEEFHORANTFTLSKFMCCHHVQVOPFE 287
 387 SGC-----TPLPMASTRARLGBELLQDFRSHQCMQMDNDMGSIMACPMKDY 436
 288 SSCINDCLRHAIVTPCPVISMVTYK-----PVADY 318
 437 IHRRWV-LVWLACLLLAALFEFL-----LLKDRRKAARS----- 472
 319 I-PLWVYGLTLIALIVGSYVILICTMWRLSGADQEKHGDSKINGILPVADLTPEPL 377
 473 --RTALLHSADGAEYELVGAALASAL-SQMPLRVAVDMSRRELSAHGALAMFHQRR 529
 378 RPRKWTIVYSADHPLVYEVVLFKFAOFLITACTEVALDLLEBVISVGVMTWSRQOE 437
 530 ILDEGVVILLES-----PAVAOCQOMIQDQVEGPHALAAMLSCVL 574
 438 MVSNSKIIILICSGTQAKKAILGMAEPAYOLRCDHW-----KPA-GDLFTAAAMNML 490
 575 PDFLOGRATGRVGVYFDGLLHPDSVSPFRVAPLESL 612
 491 PDFKRPACFGTYVVCYFSGICSERDVPDLFNITSRYPL 528

RESULT 4
 S-09-022-696-2
 Sequence 2, Application US/09022696
 Patent No. 6072037

GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin
 APPLICANT: Spriggs, Melanie
 APPLICANT: Fanslow, William
 TITLE OF INVENTION: No. 6072037el Receptor That Binds IL-17
 NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation
 STREET: 51 University Street
 CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: Apple Power Macintosh
 SOFTWARE: Microsoft Word for Apple, Version 6.0.1
 CURRENT APPLICATION DATA:
 FILING DATE:
 APPLICATION NUMBER: US/09/022.696
 CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/620,694
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USSN 08/410,535
 FILING DATE: 23 MARCH 1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia Anne
 REGISTRATION NUMBER: 34,695
 TELEPHONE: (206)587-0430
 TELEFAX: (206)
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 864 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-022-696-2

Query Match 4.6%; Score 167; DB 3; Length 864;
 Best local similarity 20.3%; Pred. No. 1,1e-08;
 Matches 105; Conservative 62; Mismatches 167; Indels 184; Gaps 25;

224 KNLN-GRONTLN-----HTDLPCLCIGVSLSPD-----SER 256
 66 KNLPPSPKNTIYINLSVSTGELVPLVHVE-WTLOTASILYLEGAEISVLQINTNER 124
 257 --VECFPRREDPGAHNRNHLRLR-VLSPG-----VMQDAPCCLPKVTLCWQAPDQ 307
 125 LCVKF-QFLSMLOHRRKRFRFSHFVDPGQGEYEVTHLPRK-----IPDG 171
 308 SPQCP-----LVPPVQKNATVNEPQDFQVLVAGHPNLCVOYSTWEKVLQACSWADSLCP 363
 172 DPNKSKITIEVPDCEDESKMKTTSCVSSGLMDPNITVETLDTQHLKVDFTLMNEST-PY 230
 364 KDMLVEMKGTGLNNTSY-----CA-----LEP-- 386
 231 Q---VLLESFSDSEHSCFEDVVKQIFAPRQEEFHORANTFTLSKFMCCHHVQVOPFE 287
 387 SGC-----TPLPMASTRARLGBELLQDFRSHQCMQMDNDMGSIMACPMKDY 436
 288 SSCINDCLRHAIVTPCPVISMVTYK-----PVADY 318
 437 IHRRWV-LVWLACLLLAALFEFL-----LLKDRRKAARS----- 472
 319 I-PLWVYGLTLIALIVGSYVILICTMWRLSGADQEKHGDSKINGILPVADLTPEPL 377
 473 --RTALLHSADGAEYELVGAALASAL-SQMPLRVAVDMSRRELSAHGALAMFHQRR 529
 378 RPRKWTIVYSADHPLVYEVVLFKFAOFLITACTEVALDLLEBVISVGVMTWSRQOE 437
 530 ILDEGVVILLES-----PAVAOCQOMIQDQVEGPHALAAMLSCVL 574
 438 MVSNSKIIILICSGTQAKKAILGMAEPAYOLRCDHW-----KPA-GDLFTAAAMNML 490
 575 PDFLOGRATGRVGVYFDGLLHPDSVSPFRVAPLESL 612
 491 PDFKRPACFGTYVVCYFSGICSERDVPDLFNITSRYPL 528

RESULT 5
 US-08-978-773-2
 Sequence 2, Application US/08978773
 Patent No. 6083906
 GENERAL INFORMATION:
 APPLICANT: Trout, Anthony
 TITLE OF INVENTION: Method of Regulating Nitric Oxide Production
 NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple PowerMacintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for PowerMacintosh, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,773

FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/052,525
FILING DATE: 27 NOVEMBER 1996
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2623-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 864 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
S-08-978-773-2

Query Match: 4.6%; Score 167; DB 3; Length 864;
Best Local Similarity 20.3%; Pred. No. 1.1e-08;
Matches 105; Conservative 62; Mismatches 167; Indels 184; Gaps 25;

224 KNTL--GPNITLN-----HTDLVPCICIOVMSLEPD-----SER 256
66 KNTLSSPKNIYINLSVSTOGELVPLVHVE-WTLOTDAISILYLEGAELSVLOANTNER 124
257 --VEFCPRREDGAGRNIMHIALRL-VISPG-----VWOLDAPCCLPKVTLCWQAPDQ 307
125 LCVKF-QELSMLOHHRKMRFSFSHFVVDPGQEVETVHHLPRP-----IPDG 171
308 SPCOP---LVPPVPQKATVNEPDQFQLVAGHPNLCVOVSTWEKYOLACSNADSLGPF 363
172 DPNHKSIIIFVPPDCEDSKMKMTTSCVSSGSLMDPNITVETLTQHLKVDFTLMNEST-PY 230
364 KDMILVEMKTGLNNTSV-----CA-----LEP-- 386
231 Q---VILSESDSENHSCFDVYKQIFAPROEERHORANVTFTLSKPHWCCHNHVOYOPFF 287
387 SGC-----TPLPMASTRARALGEELLQDFRSHOCMOLMNDNMMSLWCPMDKY 436
288 SSCINDCLRHAIVTPCPYISMTYPR-----PADY 318
437 IHRWV--LVWLACILLAALFFEL-----LLKDRKRAAGS----- 472
319 I-PLWVYGLTLILILVGSIVILYICMTWRLSGADQEKHGDSKINGILPVADLTGPPL 377
473 --FTALLHNSADGAGYELVGLASAL--SQMPLRYAVUDLMSRRELSAHGALWPHHRRR 529
378 RPRKRWIVYSDHDLVYEVVLFKFAOFLITAGCTEVALDLLEEVYISEVGVTWVSROKOE 437
530 ILDEGVVILLES-----PAAVAOCQOMLOLOTVEGPRHDLAAMLSCVL 574
438 MVSNSKIIILCSRGTOAKWKAILGMAEPAYQLRCDHM-----KPA-GDLFTAAANMIL 490
575 PDFLOGRATGATVGVYIDGLLHPDSVSPFRVAPLESL 612

Db 491 PDFKRPACFGTYVVCYFSGICSERDVPDLNITSRPL 528

RESULT 6

US-09-022-253-2
Sequence 2, Application US/09022253
Patent No. 6096305
GENERAL INFORMATION:
APPLICANT: Yao, Zhenqin
APPLICANT: Spriggs, William
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6096305el Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,253
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/620,694
FILING DATE: 21-MARCH-1996
APPLICATION NUMBER: USSN 08/538,765
FILING DATE: 7 AUGUST 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995

ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 864 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-253-2

Query Match: 4.6%; Score 167; DB 3; Length 864;
Best Local Similarity 20.3%; Pred. No. 1.1e-08;
Matches 105; Conservative 62; Mismatches 167; Indels 184; Gaps 25;

224 KNTL--GPNITLN-----HTDLVPCICIOVMSLEPD-----SER 256
66 KNTLSSPKNIYINLSVSTOGELVPLVHVE-WTLOTDAISILYLEGAELSVLOANTNER 124
257 --VEFCPRREDGAGRNIMHIALRL-VISPG-----VWOLDAPCCLPKVTLCWQAPDQ 307
125 LCVKF-QELSMLOHHRKMRFSFSHFVVDPGQEVETVHHLPRP-----IPDG 171
308 SPCOP---LVPPVPQKATVNEPDQFQLVAGHPNLCVOVSTWEKYOLACSNADSLGPF 363
172 DPNHKSIIIFVPPDCEDSKMKMTTSCVSSGSLMDPNITVETLTQHLKVDFTLMNEST-PY 230
364 KDMILVEMKTGLNNTSV-----CA-----LEP-- 386


```

231 Q---VLESFSDSENHSCFVYVKQIFAPROEEFHQRANVTTLTSKFWHCCHHHVOVQPF 287
387 SGC-----TPLESMASTRAARLGEELLQDFRSHQCMQMDNDNMSGLMACPMKY 436
288 SSCNDCLRHAVTVPCPVISNTTVPK-----PVADY 318
437 IHRWV--LWMLACLLLAALFFFL-----LKKRRKKAARS-----472
319 I-PLWVYGLITLAILLVGSVIVLICTMWRLSGADQEKHDDSKINGILPVADLTPLPL 377
473 --RTALLHSADAGCYERLYGALASAL-SQMLPVAVDMSRRELISANGALAMFHQR 529
378 RPRKWIYVSADHPLVEVYLKRAQFLITACGTEVALDLLEOVISEVGMVWSROKOE 437
530 ILQEGGVILLFS-----PAVAOCQOMLQGTVEPGRHDALAAMLSCVL 574
438 MYESNKIILCSRGTOAKKAILGNAEPVQLKCDHM-----KPA-GDLFTAAANMIL 490
575 PDFLOGRATGRVGVYFDGLLHPDSVSPFRVAPLFSL 612
491 PDFKRPACGTYVVCYFSGICSERDVPDLFNITSRYPL 528

```

RESULT 7
US-09-022-260-2
Sequence 2, Application US/09022260
Patent No. 6100235
GENERAL INFORMATION:
APPLICANT: Yao, Zhengbin
APPLICANT: Spriggs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6100235el Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: Apple Power Macintosh
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,260
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/620,694
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 864 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-260-2

```

Query Match 4.6%; Score 167; DB 3; Length 864;
Best Local Similarity 20.3%; Pred. No. 1,1e-08;
Matches 105; Conservative 62; Mismatches 167; Indels 164; Gaps 25;

```

224 KULT--GPNQNTLN-----HTDLVPLCLIQVWSLEPD-----SER 256
66 KNTLFPSPKNTIYINLVSSYOHGELVPLHVE-WTLQTDASTLYLEGAEISYLQNTNR 124
257 --VEFCPPFEDPAHNMIMIARLR-VLSFG-----WQDAPCCLCKYLCWAPQ 307
125 LCVKF-QELSMLOHHKRWKFSFHFVYVDPQOEYEVTVHLLPKP-----IPG 171
308 SPCOP---LVPPVPKQATVNEPQDFOLVAGHPNLQVQSTWEKVOLOACSMADSLGPF 363
172 DENHRSKIIFVPPCEDSKMNTTSCVSSGLMDPNITVETIDTQHLRVDFTLNMEST-PY 230
364 KDDMLVEKKTGINTSV-----CA-----LEP-- 386
231 Q---VLESFSDSENHSCFVYVKQIFAPROEEFHQRANVTTLTSKFWHCCHHHVOVQPF 287
387 SGC-----TPLESMASTRAARLGEELLQDFRSHQCMQMDNDNMSGLMACPMKY 436
288 SSCNDCLRHAVTVPCPVISNTTVPK-----PVADY 318
437 IHRWV--LWMLACLLLAALFFFL-----LKKRRKKAARS-----472
319 I-PLWVYGLITLAILLVGSVIVLICTMWRLSGADQEKHDDSKINGILPVADLTPLPL 377
473 --RTALLHSADAGCYERLYGALASAL-SQMLPVAVDMSRRELISANGALAMFHQR 529
378 RPRKWIYVSADHPLVEVYLKRAQFLITACGTEVALDLLEOVISEVGMVWSROKOE 437
530 ILQEGGVILLFS-----PAVAOCQOMLQGTVEPGRHDALAAMLSCVL 574
438 MYESNKIILCSRGTOAKKAILGNAEPVQLKCDHM-----KPA-GDLFTAAANMIL 490
575 PDFLOGRATGRVGVYFDGLLHPDSVSPFRVAPLFSL 612
491 PDFKRPACGTYVVCYFSGICSERDVPDLFNITSRYPL 528

RESULT 8
US-09-022-259-2
Sequence 2, Application US/09022259
Patent No. 6191104
GENERAL INFORMATION:
APPLICANT: Yao, Zhengbin
APPLICANT: Spriggs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6191104el Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: Apple Power Macintosh
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,259
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/620,694
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 866 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-255-10

Query Match 4.28; Score 151.5; DB 3; Length 866;
Best Local Similarity 19.28; Pred. No. 5.4e-07;
Matches 142; Conservative 73; Mismatches 206; Indels 317; Gaps 34;

55 SAPGP-----VLP-----TRQTELVLRCPQKTCALRVVYVHLAVHGHMAEPE 100
10 AVRPRLGLLLLLGLVLPAGGASRLDLHRALVCSQP---GLMCTYKNSCTCLDDSWIHPR 66
101 EAGKSELDSESRNASLQAOVYLSF---QAYPIARCALLEVOVPADLVOPGQSVGSAYF 156
67 -----NLTPSSPKDQIQHLFAHQGDLPVVAH---IEWTLQTD----- 103
157 DCFEASL-----GAEVQI-----WSTY-----KRYQKEL 181
104 -----ASTLYLEGAEISVLQNLNTERLCVFEELSKLRHHRRWRFFSHFVVDQDEYEV 159
182-----NLTPQOLPGDQVNLTLVDVSEODFSFLVLRVYD-----ALKSIMYKN 225
160 TVNHLPRPIPDG-----PNHOSKNFL-----VPDCEHARKKVTTPCMSSGSLMDPN 206
226 LTGPQNTLNTDVLPCICIQWLSLEPDSERVEFC-----PFREDPGANR 270
207 IT-----VETLEAHQLRVSEFTLMNESTHYQILLTSPFHENHSCFE 247
271 NLNHLIARLVLPQWQDAPRCCLRGKYLCKQAPRQSPQPLVPPRPQKATVNEPODF 330
248 HHHH-----PAPR-----DEEF 260
331 Q-----LVAGHPNLQVQSTWEKYOQACSWADSLGPFKDMILVEMKGLNNTSCAL 384
261 HORSNVTLLTNLKGCCR---HOVOIQ-----PF-----FSSCLND-----CL 295
385 EPSCGCTLPSSASTRAARLGEELLQDFRSHQCMQWMDNDMGSIMACPMDKYIHRRLV 444
296 RHSATVSCPEMPDT-----PEPIPDY-----MPLW-----VYWFIT 326
445 WLACILLAALEFFLLKKDKRKARGS-----RTALLH 479
327 GISLLLVGSYV---LLIYQMTWRLAPGSEKYSDDTKYTDGLPRAADLPPRLKPKWILY 384
480 SADGAGERYLVALAS--ALSOMPLRAVDLMSRRELASGALAMFHHORRILOEGGVI 538
385 SADHPLVVDVLRKQFLITACGTEVALDLLEQOASISGAVMTWVGROKEMKESNSKTI 444
539 LLESPAAVQOCQWL-----OLQTEBERP--HDAIAMLSCVLPDFLOGRATGRYGVYE 591
445 VLCSRGTRAKQALLIGRCAPVRLRCDHGKRVGDLFTFAANMMLPDKFRKRCFCGYVVCYF 504
592 DGLLHPDSVSPFRVAPLFSLPQLPAP-----LDALOGGCGSTSGRPRADVERVYTO-- 643
505 SEVSCDGDVDPDLFAAPRYPLMDREEVYFRIDQLEMFOPG-----RHHRVGELS 554
644 -----ALRSALD 650
555 GDMYLRSPGGRQLRAALD 572

RESULT 12

US-09-022-696-10

Sequence 10, Application US/09022696

Patent No. 6072037

GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin

APPLICANT: Spriggs, Melanie

APPLICANT: Fanslow, William

TITLE OF INVENTION: NO. 6072037e1 Receptor That Binds IL-17

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

City: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple Power Macintosh

OPERATING SYSTEM: Apple Operating System 7.5.5

SOFTWARE: Microsoft Word for Apple, Version 6.0.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/022,696

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/620,694

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/410,535

FILING DATE: 23 MARCH 1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,695

REFERENCE/DOCKET NUMBER: 2617-B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)587-0430

TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 866 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-022-696-10

Query Match 4.28; Score 151.5; DB 3; Length 866;

Best Local Similarity 19.28; Pred. No. 5.4e-07;

Matches 142; Conservative 73; Mismatches 206; Indels 317; Gaps 34;

55 SAPGP-----VLP-----TRQTELVLRCPQKTCALRVVYVHLAVHGHMAEPE 100
10 AVRPRLGLLLLLGLVLPAGGASRLDLHRALVCSQP---GLMCTYKNSCTCLDDSWIHPR 66
101 EAGKSELDSESRNASLQAOVYLSF---QAYPIARCALLEVOVPADLVOPGQSVGSAYF 156
67 -----NLTPSSPKDQIQHLFAHQGDLPVVAH---IEWTLQTD----- 103
157 DCFEASL-----GAEVQI-----WSTY-----KRYQKEL 181
104 -----ASTLYLEGAEISVLQNLNTERLCVFEELSKLRHHRRWRFFSHFVVDQDEYEV 159
182-----NLTPQOLPGDQVNLTLVDVSEODFSFLVLRVYD-----ALKSIMYKN 225
160 TVNHLPRPIPDG-----PNHOSKNFL-----VPDCEHARKKVTTPCMSSGSLMDPN 206
226 LTGPQNTLNTDVLPCICIQWLSLEPDSERVEFC-----PFREDPGANR 270
207 IT-----VETLEAHQLRVSEFTLMNESTHYQILLTSPFHENHSCFE 247

QY 271 NLMTIARLVSPGVWQDAPCCLPKVTLCWQAPDQSPQPLVPVPOKNATVNEPODF 330
 DB 248 HHHHI-----PAPR-----PEEF 260
 QY 331 Q-----LVAGHPNLCVQVSTWEKVQIACSWADSLGPFKDMILVEKMTGLNNTSVCA 384
 DB 261 HORSNVTTLRLNKGCCR-----HOVOIQ-----FSSCLND-----CL 295
 QY 385 EPGSGTLPMSASTRAARLGEELLQDFRSHOCQMLMNDNMGSLMACPMCKYIHRRLV 444
 DB 296 RHSAIVSCPEMDT-----PEPIPDY-----MPLM-----VYWFIT 326
 QY 445 WLACLLAALAEFFELLKKDRKARGS-----RTALLH 479
 DB 327 GISILLVGSVI-----LLIYCMTRILAGPSEKYSDDTKYTDGLPRADLIIPPLKPRKWTIY 384
 QY 480 SADGAGIERLVGALAS-ALSQMPLRAVAVDLSRRELSAHGALAMFHQRRILOEGGVI 538
 DB 385 SADHPLYVDVVLKFAQFLITACGTEVALDLLEQALISEAGVMTWVGROKQEWESNSKI 444
 QY 539 LLESPAANAQCCQWL-----QLOTVEGP-HDALAAMLSCVLPDFLOGRATGRVGYVF 591
 DB 445 VLCSRGTRAKWQALLGKAPVRLKCDHGKPYGDLFTAAAMNMLPDKRPACFGTYVVCYF 504
 QY 592 DGLLHPDVSPPRVAPLPSLPQLPAF-----LDALQGGCSTSGRPRADRYERTQ-- 643
 DB 505 SEVSCGDGVDPDLFGAARPRPLMDRFEERYFRIDLEMFQPG-----RMHRVGEIS 554
 QY 644 -----ALRSALD 650
 DB 555 GDNYLRSPPGROLRAALD 572

RESULT 13

IS-08-978-773-4
 Sequence 4, Application US/08978773
 Patent No. 6083906

GENERAL INFORMATION:

APPLICANT: TROUT, Anthony
 TITLE OF INVENTION: Method of Regulating Nitric Oxide Production
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: Apple Operating System 7.5.5

SOFTWARE: Microsoft Word for PowerMacintosh, Version 6.0.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/978,773

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 60/052,525

FILING DATE: 27 NOVEMBER 1996

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,693

REFERENCE/DOCKET NUMBER: 2623-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)587-0430

TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 866 amino acids

TYPE: amino acid

; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-978-773-4

Query Match 4.28; Score 151.5; DB 3; Length 866;
 Best Local Similarity 19.28; Pred. No. 5.4e-07;
 Matches 142; Conservative 73; Mismatches 206; Indels 317; Gaps 34;

QY 55 SAPCP-----VYVP-----TRLOTELVLRQCQKTCALRRVYVHLAVGHNAEPE 100
 DB 10 AVPGPLGLILLGLLVLAAPGASLRLDHRALVCSOP-----GLNCTVKNSTCLDSDWTHPR 66
 QY 101 EAGKSDSELOESRNASLOAQVLSF-----QAVPIARCALLEVOVPADLPQGSVGSASF 156
 DB 67 -----NLTPSSPKDQIQHFAHTQGDLEFVAH-----IEMTLQTD----- 103
 QY 157 DCREASL-----GAEVQI-----WSYT-----KPRYQEL 181
 DB 104 -----ASILYLEGAEISVLIQNTNRLCYRFEFLSKIRHHRRMRFTFSHFVVDPDQYEY 159
 QY 182 -----NLQOLPDGDNVLLTLDVSEBODFSFLYLKRVPD-----ALKSILMYKN 225
 DB 160 TYHHLKPLPDGD-----PNHOSKNFL-----VPDEHARMKYVTPCMSGSLMDPN 206
 QY 226 LKGPQNTLTHNTDLVPCLCIQVMSLEPDSERYEFC-----PREDDGAHR 270
 DB 207 IT-----VETLEAHQLRVSTFLMNESTHYQLILSFPHENHSCEE 247
 QY 271 NLMTIARLVSPGVWQDAPCCLPKVTLCWQAPDQSPQPLVPVPOKNATVNEPODF 330
 DB 248 HHHHI-----PAPR-----PEEF 260
 QY 331 Q-----LVAGHPNLCVQVSTWEKVQIACSWADSLGPFKDMILVEKMTGLNNTSVCA 384
 DB 261 HORSNVTTLRLNKGCCR-----HOVOIQ-----FSSCLND-----CL 295
 QY 385 EPGSGTLPMSASTRAARLGEELLQDFRSHOCQMLMNDNMGSLMACPMCKYIHRRLV 444
 DB 296 RHSAIVSCPEMDT-----PEPIPDY-----MPLM-----VYWFIT 326
 QY 445 WLACLLAALAEFFELLKKDRKARGS-----RTALLH 479
 DB 327 GISILLVGSVI-----LLIYCMTRILAGPSEKYSDDTKYTDGLPRADLIIPPLKPRKWTIY 384
 QY 480 SADGAGIERLVGALAS-ALSQMPLRAVAVDLSRRELSAHGALAMFHQRRILOEGGVI 538
 DB 385 SADHPLYVDVVLKFAQFLITACGTEVALDLLEQALISEAGVMTWVGROKQEWESNSKI 444
 QY 539 LLESPAANAQCCQWL-----QLOTVEGP-HDALAAMLSCVLPDFLOGRATGRVGYVF 591
 DB 445 VLCSRGTRAKWQALLGKAPVRLKCDHGKPYGDLFTAAAMNMLPDKRPACFGTYVVCYF 504
 QY 592 DGLLHPDVSPPRVAPLPSLPQLPAF-----LDALQGGCSTSGRPRADRYERTQ-- 643
 DB 505 SEVSCGDGVDPDLFGAARPRPLMDRFEERYFRIDLEMFQPG-----RMHRVGEIS 554
 QY 644 -----ALRSALD 650
 DB 555 GDNYLRSPPGROLRAALD 572

RESULT 14

US-09-022-253-10

Sequence 10, Application US/09022253

Patent No. 6096305

GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin

APPLICANT: Spriggs, William

TITLE OF INVENTION: No. 6096305e1 Receptor That Binds IL-17

CORRESPONDENCE ADDRESS:

Sun Sep 29 09:32:43 2002

us-09-899-471-2.rai

Page 11

Query Match	4.28;	Score 151.5;	DB 3;	Length 866;
Best Local Similarity	19.28;	Pred. NO. 5.4e-07;		
Matches 142;	Conservative 73;	Mismatches 206;	Indels 317;	Gaps 34;

[illegible]

March completed: September 28, 2002, 19:40:27
 Time: 6270 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

protein - protein search, using sw model

Search time 72.7 Seconds
(without alignments)
1603.832 Million cell updates/sec

US-09-899-471-2

Sequence: 1 MPVSMFLSLALGRNPVVS.....SSAPGCCCEMDLPCPTTLE 674

Gapop 10.0 , Gapext 0.5

562222 segs, 172994929 residues

562222

Minimum DB seq length: 0
Minimum DB seq length: 2000000000

Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

- 1: SP:archaea:*
- 2: SP:bacteria:*
- 3: SP:fungi:*
- 4: SP:human:*
- 5: SP:invertebrate:*
- 6: SP:mammal:*
- 7: SP:mhc:*
- 8: SP:organelle:*
- 9: SP:phage:*
- 10: SP:plant:*
- 11: SP:rodent:*
- 12: SP:virus:*
- 13: SP:vertebrate:*
- 14: SP:unclassified:*
- 15: SP:rvirus:*
- 16: SP:bacteriaph:*
- 17: SP:archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Rank	Score	Query	Match	Length	DB	ID	Description
1	2471	68.5	567	11	099J43	099J43	mus musculus
2	1749.5	48.5	538	4	09BR97	09BR97	homo sapien
3	167	4.6	864	11	060943	060943	mus musculus
4	151.5	4.2	866	4	043844	043844	homo sapien
5	151.5	4.2	866	4	096F46	096F46	homo sapien
6	114.5	3.2	1013	16	053499	053499	mycobacteri
7	110.5	3.1	2012	4	09BX84	09BX84	homo sapien
8	110	3.1	2910	11	055225	055225	mus musculus
9	108.5	3.0	361	3	060150	060150	schizosacch
10	108.5	3.0	478	16	09RY89	09RY89	homo sapien
11	106.5	3.0	589	10	094G03	094G03	flaviera pr
12	106.5	3.0	1248	4	090J61	090J61	homo sapien
13	106	2.9	756	10	093XR9	093XR9	bruguliera g
14	106	2.9	1693	2	09REL2	09REL2	agrobacteri
15	105.5	2.9	698	16	09PBH5	09PBH5	xytelia fas

17	105	2.9	1767	5	09W4F5	09W4F5	drosophila
18	105	2.9	1767	5	024495	024495	drosophila
19	104.5	2.9	549	10	09FV22	09FV22	cinnamomum
20	104.5	2.9	581	10	094BW5	094BW5	cinnamomum
21	104.5	2.9	864	4	043297	043297	homo sapien
22	104.5	2.9	1237	4	09UNC9	09UNC9	homo sapien
23	104.5	2.9	1248	4	09UBZ5	09UBZ5	homo sapien
24	104.5	2.9	3021	12	081258	081258	hepatitis c
25	103.5	2.9	383	5	09VIT5	09VIT5	hepatitis c
26	103.5	2.9	936	13	09BS71	09BS71	hepatitis c
27	103	2.9	851	5	09U2S3	09U2S3	caenorhabdi
28	103	2.9	906	16	09HX92	09HX92	pseudomonas
29	102.5	2.8	758	4	09NKR9	09NKR9	homo sapien
30	102.5	2.8	1239	13	09DEF4	09DEF4	homo sapien
31	102.5	2.8	1418	13	09BSW6	09BSW6	xenopus lae
32	102	2.8	509	11	008907	008907	xenopus lae
33	102	2.8	758	6	09NDH2	09NDH2	sus scrofa
34	102	2.8	955	4	096DN2	096DN2	homo sapien
35	102	2.8	3021	12	068870	068870	hepatitis c
36	101.5	2.8	610	4	014775	014775	homo sapien
37	100.5	2.8	2270	4	09NSJ3	09NSJ3	homo sapien
38	100	2.8	3013	12	092530	092530	hepatitis c
39	100	2.8	3019	12	092529	092529	hepatitis c
40	99.5	2.8	589	10	094G02	094G02	flaviera pr
41	99.5	2.8	1114	11	09JKW7	09JKW7	mus musculus
42	99	2.7	932	11	09D2V7	09D2V7	mus musculus
43	99	2.7	1002	11	091VB4	091VB4	mus musculus
44	99	2.7	1617	13	073905	073905	gallus gall
45	99	2.7	2308	5	09VJD9	09VJD9	drosophila

ALIGNMENTS

RESULT	1	PRELIMINARY	PRT	567 AA.
ID	099J43	099J43		
AC	099J43	099J43		
DT	01-JUN-2001	(TREMBLrel. 17, Created)		
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)		
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)		
DE	HYPOTHETICAL 62.8 KDA PROTEIN.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RA	Strausberg R;			
RL	Submitted (Mar-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: BC004759; AAH04759.1;			
KW	Hypothetical protein.			
SO	SEQUENCE 567 AA; 62798 MW; C1AAB79E2006BID CRC64;			
Query Match	68.5%; Score 2471; DB 11; Length 567;			
Best Local Similarity	94.7%; Pred. No. 7.3e-247			
Matches	463; Conservative 0; Mismatch 2; Indels 24; Gaps 1;			
QY	1	MPVSMFLSLALGRNPVVSLERLMEPODTRCSIGLSCHLMDGVICLPGLSAPGPV 60		
DB	1	MPVSMFLSLALGRNPVVSLERLMEPODTRCSIGLSCHLMDGVICLPGLSAPGPV 60		
QY	61	LVPTRLQTELVLRCPKTDALRVVVVHLAVHGMABPEEAGKSDSELRNASTLQAO 120		
DB	61	LVPTRLQTELVLRCPKTDALRVVVVHLAVHGMABPEEAGKSDSELRNASTLQAO 120		
QY	121	VVLSQAVYFIARCALLLEVQVADLVQGVSGSAVFDCFEASLGAEOVTWSTKPRYQKE 180		
DB	121	VVLSQAVYFIARCALLLEVQVADLVQGVSGSAVFDCFEASLGAEOVTWSTKPRYQKE 180		
QY	181	LVLTQQLP-----DGDVLTITDVSEEDPSFLILYLRPYD 216		
DB	181	LVLTQQLP-----DGDVLTITDVSEEDPSFLILYLRPYD 216		

181 LNLTOOLPDCRGLEVDSDIOSCWYLPWLNSTDDNVLTLTLDVSEEDSEFLYLPRVPD 240
217 AKSLMYKMLTGPONITLNTDLPVCLIOVWSLEPDSERVEFCFREDPGAHRLMHLA 276
241 AKSLMYKMLTGPONITLNTDLPVCLIOVWSLEPDSERVEFCFREDPGAHRLMHLA 300
277 RLRLVSPGWMDAPCCIPKVTLCQAPDPSPCOPLVPPVPOKNAVNEPQFOLVAGH 336
301 RLRLVSPGWMDAPCCIPKVTLCQAPDPSPCOPLVPPVPOKNAVNEPQFOLVAGH 360
337 PNLGVSTWKEVQVLOACSMADSLGPPKDDMLLVEMKTGLNNTSVCALEPSCGTPSPMA 396
361 PNLGVSTWKEVQVLOACSMADSLGPPKDDMLLVEMKTGLNNTSVCALEPSCGTPSPMA 420
397 STRARLGEELLQDFRSHQCMQMLNDNMGSLMACPMKTYHRRVYLVLAACLLAAALF 456
421 STRARLGEELLQDFRSHQCMQMLNDNMGSLMACPMKTYHRRVYLVLAACLLAAALF 480
457 FELLKKDR 465
481 FELLKKDR 489

ESLUT 2
PRELIMINARY: PRT; 538 AA.
09BR97
01-JUN-2001 (TREMBLrel. 17, Created)
01-JUN-2001 (TREMBLrel. 17, last sequence update)
01-DEC-2001 (TREMBLrel. 19, last annotation update)
HYPOTHEICAL. 59.1 KDA PROTEIN.
Homo sapiens (Human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_Taxid=9606;
(1)
SEQUENCE FROM N.A.
TISSUE-ENOMERIAL ADENOCARCINOMA;
Strausberg R.;
Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
EMBL; BC006411; AA06411.1;
Hypothetical protein.
SEQUENCE 538 AA; 59127 MW; 65E6344DA6A5AFD2 CRC64;

Query Match 48.5%; Score 1749.5; DB 4; Length 538;
Best Local Similarity 65.2%; Pred. No. 6.5e-149;
Matches 339; Conservative 51; Mismatches 107; Indels 23; Gaps 5;
1 MPVSWFLSLALGRNPNVYSLERLMEPQDARCSIGLSCHLMDGVDLPGSLQSPAGPV 60
1 MPVSWFLSLALGRNPNVYSLERLMEPQDARCSIGLSCHLMDGVDLPGSLQSPAGPV 60
61 LVPRLQTELVLRCPQKTDCAKRVVYVHLAVHGMWAEPEE---AGKSDSELQESRNAS 116
61 LVPRLQTELVLRCPQKTDCAKRVVYVHLAVHGMWAEPEE---AGKSDSELQESRNAS 116
61 LAPRLQTELVLRCPQKTDCAKRVVYVHLAVHGMWAEPEE---AGKSDSELQESRNAS 120
117 LQAVVYLSQAPRIACALLEVQPADLVQPGSVGSAVFCFPEASLAEVQIWSYTKR 176
121 LQAVVYLSQAPRIACALLEVQPADLVQPGSVGSAVFCFPEASLAEVQIWSYTKR 180
177 YOKELNLTQOOLP-----DGDNVLTLTLDVSEEDSEFLYLPRVPDALKSLMYKMLT 227
181 YOKELNLTQOOLP-----DGDNVLTLTLDVSEEDSEFLYLPRVPDALKSLMYKMLT 240
228 GPONITLNTDLPVCLIOVWSLEPDSERVEFCFREDPGAHRLMHLA 287
241 GPONITLNTDLPVCLIOVWSLEPDSERVEFCFREDPGAHRLMHLA 300
288 LDPACCLGKVTLCQAPDPSPCOPLVPPVPOKNAVNEPQFOLVAGH 347
301 LDPACCLGKVTLCQAPDPSPCOPLVPPVPOKNAVNEPQFOLVAGH 360

QY 348 KYOLACSMADSLGPPKDDMLLVEMKTGLNNTSVCALEPSCGTPSPMASTRARLGEEL 407
Db 361 KYOLACSMADSLGPPKDDMLLVEMKTGLNNTSVCALEPSCGTPSPMASTRARLGEEL 420
QY 408 LQDFRSHQCMQMLNDNMGSLMACPMKTYHRRVYLVLAACLLAAALFELLKKDR 467
Db 421 LQDFRSHQCMQMLNDNMGSLMACPMKTYHRRVYLVLAACLLAAALFELLKKDR 477
QY 468 AARGSFALLHSDAGYERLV-----GALASALSCMP 501
Db 478 AARGSFALLHSDAGYERLV-----GALASALSCMP 516

RESULT 3
ID 060943 PRELIMINARY: PRT; 864 AA.
AC 060943;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE INTERLEUKIN 17 RECEPTOR.
GN IL17R.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_Taxid=10090;
RX (1)
SEQUENCE FROM N.A.
TISSUE-THYMOMA EL4.
MEDLINE=96111968; PubMed=8777726;
RA Yeo Z., Fanslow W.C., Seidlin M.F., Rousseau A.M., Painter S.L.,
RT Comeau M.R., Cohen J.I., Spitz M.K.,
RT "Herpesvirus Saimiri encodes a new cytokine, IL-17, which binds to a
RL novel cytokine receptor."
RL Immunity 3:811-821(1995).
DR EMBL; U31993; AAC52357.1;
DR MGI; MGI:107399; 11117F.
KW Receptor.
SO SEQUENCE 864 AA; 97807 MW; 343FD51AA687DA31 CRC64;

Query Match 4.6%; Score 167; DB 11; Length 864;
Best Local Similarity 20.3%; Pred. No. 4.3e-06;
Matches 105; Conservative 62; Mismatches 167; Indels 184; Gaps 25;
QY 224 KNLTL--GPONITLNT-----HTDLPVCLIOVWSLEPD-----SER 256
Db 66 KNLTLSPSPNITNIVSVSSQHGELVPLVHVE-WTQTDASILYEGAEISVLQNTNER 124
QY 257 --VERCFREDPGAHRLMHLA 307
Db 125 LCVKF-QFLSMQNHRRKMFSSHFVVDGQEVYVNHLPKP-----IPDG 171
QY 308 SPCOP---LVPVPOKNAVNEPQDQFOLVAGHNPVLCVQVSTWKEVQVLOACSMADSLGPF 363
Db 172 DPNHKSILTFVDECSKAKMTTSCVSSGSLMDPNTVEITDQHLRVDFLTMINST-PY 230
QY 364 KDDMLLVEMKTGLNNTSV-----CA-----LEP-- 386
Db 231 Q---VLLSEFSSEHNSCDFVQKIFARQOEFGQNAVTFELTSFHCCHNHVQVPEF 287
QY 387 SGC-----TLPASASTRAARLGEELLQDFRSHQCMQMLNDNMGSLMACPMKTY 436
Db 288 SSCLNDCLRHAVTVPCVVISNTVTPK-----PVADY 318
QY 437 IHRNV--LVNLACLLAAALFFEL-----LKKDRRKARG-- 472
Db 319 I-PLVWYGLITLAILLVSVYVLIICMTWRLSGADDEKDDSKINILTPADITPPPL 377
QY 473 --RTALLHSADGAGYERLVGALASAL--SOMPLKVAVDLWSRRELTAHAGALMFFHQR 529
Db 378 RPRKWIYVYSAHPIVVEVYLVKFAQELITACGTEVADLLEQVYISEGVMTWVSROKE 437

530 ILDEGVVILLES-----PAVAOCQOMLOQTEVEPGPHDLAAMLSCVL 574
 438 MESSKIIILCSRGTOAKKAILGWAEPVQLRCDHM-----KPA-GDLFTAAAMNML 490
 575 PDFLOGRATGRVGVYFDDLLHPDSVSPFRVAPLFSL 612
 491 PDFKRPACRGTYVVCFSGICSEBDVDFLNTSRYP 528

RESULT 4

ID 043844 PRELIMINARY: PRT: 866 AA.
 043844; 01-JUN-1998 (TREMBlrel. 06, Created)
 01-JUN-1998 (TREMBlrel. 06, last sequence update)
 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 IL-17 RECEPTOR.
 Homo sapiens (human).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NCBI_Taxid=9606;
 [1]
 SEQUENCE FROM N.A.
 MEDLINE=98035683; Pubmed=9367539;
 Yao Z., Spriggs M.K., Dery J.M.J., Strockbine L., Park L.S.,
 Vandenbos T., Zappone J., Painter S.L., Amilage R.J.;
 "Molecular characterization of the human interleukin (IL)-17
 receptor.";
 Cytokine 9:794-800(1997).
 EMBL: U58917; AAB9730.1;
 RECEPTOR.
 SEQUENCE 866 AA; 96122 MW; 88AF626A83FE70 CRC64;

Query Match 4.2%; Score 151.5; DB 4; Length 866;
 Best Local Similarity 19.2%; Pred. No. 0.00011;
 Matches 142; Conservative 73; Mismatches 206; Indels 317; Gaps 34;

55 SAPGP-----VLVP-----TRLOTELVRCPQKTCALRRVVVHLAVHGMARE 100
 10 AVPGPLGLLLGLLVLPAGASLRDLHRLVCSGP---GLNCTVANSCTLDSDWIHR 66
 101 EAGKSDSELRNNAQAVLSF---QAVPIARCALLEVOVPALVOPGSGSAVF 156
 67 -----NLTPSSPKDLOQLFAHTQOGDLFPVAH---TEMTLOTD----- 103
 157 DCFEASL---GAEQVQI-----MSYT-----KRYOKEL 181
 104 -----ASTILYEGAEISVQLNTNERLCVREFLSKLRHHRHRRFTSHFVVDQEEYV 159
 182 -----NLTOOLPGDGNVLLDVSEBDFSLYLRYPD-----ALKSMTKN 225
 160 TVNHLRPPIPGD-----PNHOSKNEL-----VPDEHARKYTTPCMGSSGLMDPN 206
 226 LTGPQNTLNTDLPVPCICIOVWSLEPDSERVEFC-----PEREDPGAHR 270
 207 IT-----VETLEAHQLRVSTFLMNESTHYOILLTSPHMHNSCFE 247
 271 NLMIHARLVLSPGVWOLADAPCCLPKGYTLCQWAPDQSCQPLVPVPQKNATVNEPDF 330
 248 HMHNI-----PAPR-----PEEF 260
 331 Q-----LVAGHPNLVGVSTWEKVOLOACSWADSLGPRKDMMLVEMKGLNNTSVCL 384
 261 HOSNVTLETLNLKGCGR---HOVOIO-----FSSCLND---CL 295
 385 EPSGCTPLPSMASTRARLGEELLQDFRSHOCMOLWNDNMGSLMACPMKYIHRWVLY 444
 296 RHSAIVSCPEPRPT-----PEPIPDY-----MPLW-----VYMET 326
 445 WLACLLIAAALFFELLKKRRKARGS-----RTALLLH 479
 327 GISILLVGSYI---LLIYCMTRLAGPGSEKYSDDRYTGDLPRAADLIPPLKPKKWMITY 384

480 SADGAGYERLYGALAS-ALSQMPLRYAVDUMSRRELSAHGALAMFHHRRILOEGVYI 538
 385 SADHPLVDVYVAKFQFLTLTACGTEVALDLEBOAISSEAVMTWVGROKOBVESNSKII 444
 539 LLFSEPAVAOCQOML-----OLOTEVEPGP-HDALAAMLSCVLPDFLOGRATGRVGVY 591
 445 VLCSRGTRAKNALLGRGAPVRLRCHGKRGVGLFTRAAAMNMLPDFKRPACRGTYVVCFS 504
 592 DGLHPDSVSPFRVAPLFSPLTPQPAF-----LDALOGCSTSAGRPADRVERVTO-- 643
 505 SEVSCDGVPPDLFGAARPRPLMDRFEEYFRIDLEMFQPG-----RMHRVGEIS 554
 644 -----ALRSALD 650
 555 GDNYLRSPGROLRAALD 572

RESULT 5

ID 096F46 PRELIMINARY: PRT: 866 AA.
 AC 096F46;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE SIMILAR TO INTERLEUKIN 17 RECEPTOR.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=UTERUS, AND LEIOMYOSARCOMA;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC011624; AAL11624.1;
 KW Receptor.
 SQ SEQUENCE 866 AA; 96131 MW; 28330BED2303B0C9 CRC64;

Query Match 4.2%; Score 151.5; DB 4; Length 866;
 Best Local Similarity 19.2%; Pred. No. 0.00011;
 Matches 142; Conservative 73; Mismatches 206; Indels 317; Gaps 34;

55 SAPGP-----VLVP-----TRLOTELVRCPQKTCALRRVVVHLAVHGMARE 100
 10 AVPGPLGLLLGLLVLPAGASLRDLHRLVCSGP---GLNCTVANSCTLDSDWIHR 66
 101 EAGKSDSELRNNAQAVLSF---QAVPIARCALLEVOVPALVOPGSGSAVF 156
 67 -----NLTPSSPKDLOQLFAHTQOGDLFPVAH---TEMTLOTD----- 103
 157 DCFEASL---GAEQVQI-----MSYT-----KRYOKEL 181
 104 -----ASTILYEGAEISVQLNTNERLCVREFLSKLRHHRHRRFTSHFVVDQEEYV 159
 182 -----NLTOOLPGDGNVLLDVSEBDFSLYLRYPD-----ALKSMTKN 225
 160 TVNHLRPPIPGD-----PNHOSKNEL-----VPDEHARKYTTPCMGSSGLMDPN 206
 226 LTGPQNTLNTDLPVPCICIOVWSLEPDSERVEFC-----PEREDPGAHR 270
 207 IT-----VETLEAHQLRVSTFLMNESTHYOILLTSPHMHNSCFE 247
 271 NLMIHARLVLSPGVWOLADAPCCLPKGYTLCQWAPDQSCQPLVPVPQKNATVNEPDF 330
 248 HMHNI-----PAPR-----PEEF 260
 331 Q-----LVAGHPNLVGVSTWEKVOLOACSWADSLGPRKDMMLVEMKGLNNTSVCL 384
 261 HOSNVTLETLNLKGCGR---HOVOIO-----FSSCLND---CL 295
 385 EPSGCTPLPSMASTRARLGEELLQDFRSHOCMOLWNDNMGSLMACPMKYIHRWVLY 444

296 RHSAWVSCPEMPDT-----DEPIDY-----MPLW-----VYMFIT 326
 445 WACLLLAALFFLLKDKRKAARG-----RTALLH 479
 327 GISILLVGVY--LLIYCMTRLAGPGSEKSDPKRYTDGLPVADLLPPLKPKRWIIY 384
 480 SAGAGYERLVGALAS-ALSGMPLRVAVDIMSRRELHANGALAMFHQRRILOEGGVY 538
 385 SAHPLVVDVYLAKRQFLIACGTEVALDLEQALISEAGVMTWGRQKQEWESNSKII 444
 539 LIESPAVAOCQOML-----OLQTEBERP--HDALAAMLSCVLPDFLOGRATRYGVYF 591
 445 VLCSRGTRAKQALLNGAPVRLRCDHGRKPVGDLFTAAANNMLPDRKRCFGTYVYCF 504
 592 DGLLHPDSVSPFVAPLFSLEPTQLPAF-----IDALQGGCSTSGRPADEVRYTQ-- 643
 505 SEVSCDGDVDFLFAAARYPLMDREFEEYFRIODEMFGPG-----RMHVRGELS 554
 644 -----ALRSALD 650
 555 GDNVYLSPGGRQRLRALD 572

RESULT 6 PRELIMINARY; PRT: 1013 AA.

053499: 01-JUN-1998 (Tremblrel. 06, Created)
 01-JUN-1998 (Tremblrel. 06, Last sequence update)
 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 HELICASE.
 HELZ OR RV2101 OR MTVO20 01.
 Mycobacterium tuberculosis.
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=1773;

SEQUENCE FROM N.A.
 STRAIN-H37RV;
 MEDLINE-98295987; PubMed-9634230;
 Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holtroyd S.,
 Horsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 Sultston J.E., Taylor K., Whitehead S., Barrell B.G.;
 "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence."
 Nature 393:537-544(1998).
 -1. SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.
 EMBL; AL021924; CAI17284.1;
 Tuberculist; RV2101;
 InterPro; IPR001410; DEAD.
 InterPro; IPR001650; Helicase_C.
 InterPro; IPR000330; SNF2_N.
 Pfam; PF00271; helicase_C.1.
 Pfam; PF00176; SNF2_N.1.
 SMART; SM00487; DEXDC; 1.
 SMART; SM00490; HELIC; 1.
 ATP-binding; Complete proteome; Helicase.
 SEQUENCE 1013 AA; 111029 MW; 601FDF1DB5CABEF1 CRC64;

Query Match 3.28; Score 114.5; DB 16; Length 1013;
 Best Local Similarity 20.48; Pred. No. 0.28; Mismatches 232; Indels 251; Gaps 33;
 Matches 143; Conservative .75;

90 LAVHGWAE--EAGKSDSELESRNASLQAVLSFQAVPIARCALLEVVPADL--- 144
 2 LVLHGWSNSGGRMLAEADSDLL---VKSPQALRSARPHFPA-----APADLLNG 49

QY 145 VQSGSVSAVFDCFEASLGEVQIMSTYKPKYQKELNITQQLPDGDNVLLTLDVSEED 204
 DB 50 IHGK-----PAAVILLPLSRAPLD 71
 QY 205 FSPFLYLRPVVDALKSLWKMLTGPONITLNHD-LVPLCLQVMSLEPSEVERPEPCFR 263
 DB 72 SPELIRLAPRAA-----RTDPLAMTVPVVDLDPTPALAALFQOP 113
 QY 264 EDGAGHNLMIHARLVLPSPGWQDAPCLPGV-----TLGV-----QAPDS 308
 DB 114 PDVRYGASVDYLAEIAYARELVE-----KGVLPQLRDTGHAACRPVYLGHDV 166
 QY 309 PCQPLV---PPVQKNATVNEPQDFOLVA-----GHPNLQV 342
 DB 167 AMSLVSAMPVCAEYVGHDPHELATSLDAMVDAVPAALSMDLLPRRGRSKHRA 226
 QY 343 VSTW-----EKVQLOCS-----WAD-----SLGPKDMILVEAKTGLNNTS 380
 DB 227 VEAMLDALTCPPGRFDEPDELALAEALRPWDVIGITGVPARATFRLESEVETENET- 285
 QY 381 VCALPEPGCTPLPMSASTRAARLGEELLQDFRSHOC-----QLMNDNMGSLMACPMKY 436
 DB 286 -----PAG-----SLMRL-ELLQSTQDPSILVPAEQAMND--GSL----- 319
 QY 437 IHRWV---LWVLAALLAALFFLLKDKRKAARGSRALLHSDAGYERLVGA 492
 DB 320 --RRWLDROPQELLTELGRASRIPELV-----PALFTAPCSGLELDADGA--YRPLSG 369
 QY 493 LASALSGMPLRVAVDIMSRRELHANGALAMFHQRRILOEGGVYLLRSP--AAVAQCO 550
 DB 370 TAAVLDEAGFGVILPSW-----W---DRRRL--GLVLSAVTPVGVYKAS 411
 QY 551 OMLOLOVE-----PGPHDALAAMLSCLP-----DFLOG 580
 DB 412 KFERQELVEERWELAVGDDPLSEEEIALTEKSPILIRRGVVALDTQMRGLETLE 471
 QY 581 RATGRYGVYFDGLL--HPDSVPFVAPLFSLEPTQLPAFIDALQGGCSTSGRPADEV 638
 DB 472 KPTGRKTTAEILALAAHPDQVDVPLEVTAV-----RADGMIGDLILAAGAAASLQPLDP 526
 QY 639 ERYTQALRSALDSCTS-----SSEARGCCEMDGPCITLLE 674
 DB 527 DCFATILRPYQKGLAMWLAFSLSSLGSLGSLADMDGLKTVQ 567

RESULT 7 PRELIMINARY; PRT: 2012 AA.

09BX84;
 AC 09BX84; 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
 DE CHANNEL-KINASE 2.
 GN CHAK2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-KIDNEY;
 RX MEDLINE-99147092; PubMed-10021370;
 RA Ryzanov A.G., Pavur K.S., Dorovkov M.V.;
 RT "Alpha-kinases: a new class of protein kinases with a novel catalytic
 domain."
 RL Curr. Biol. 9:R43-R45(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-KIDNEY;
 RA Ryzanova L.V., Pavur K.S., Petrov A.N., Dorovkov M.V., Ryzanov A.G.;
 RT "Novel type of signaling molecules: protein kinases covalently linked
 to ion channels."
 to ion channels."

MOI. Biol. 0-0-0(2001).
 DR EMBL: AF350881; AAK31202.1;
 DR InterPro: IPR000636; Cation_chan_non_119.
 DR InterPro: IPR002111; Cat_channel_TPL.
 DR Pfam: PF00520; Ion_trans_1.
 DR Kinase.
 KW SEQUENCE 2012 AA; 230493 MW; 8AD583235D8080AF CRC64;

Query Match 3.1%; Score 110.5; DB 4; Length 2012;
 Best Local Similarity 19.6%; Pred. No. 1.7;
 Matches 138; Conservative 84; Mismatches 269; Indels 213; Gaps 30;

1 MPVSFLLSLAGRPVYVSLERLMPDPTAR-----CSIGLS 38
 1268 MPSS-LKRLSLAGRRPPRYORCALLETITNSKREATVNRNDQERQETSSIVSGVSPNRQ 1326
 39 CHLMDGVLCIPGSIQASAPGVLPTRLCQTELYLCRQ-----KIDCALRRVYVYHLA-- 91
 1327 AHSKYGQFLVPSNLKR-----VPSAEVLPILSRSPVDVLTATEDQIQTEVYHLTGO 1380
 92 --VHGMA--EPEBAGKSDSEIQESRNASIQAOVLSFOAYPIARCALLEYQVADLWQ 146
 1381 TPVVSQMASVDEPKKEKHEPIAHLIDGQDAEOVLPITLSCTPEPMTSSPLS---QAKIWQ 1437
 147 PGQSVGSAVPDCEASLGAEOIWSYTK-----PRYQKELNTLOQLPDCGN 192
 1438 TGGGVYNNAF-----SEGDETVGSIKKKWQTCLPSTCDSSRSRQKQAKQDSSLSGN 1492
 193 VLLTLDVEEDDFSLIYLRPPDAKLSMTYKMLTGQNTLTNTDVLPCICIQWLSLP 252
 1493 STRAQSECESEVG--PWLP--NTSFWINPLRRYRPPARSHS-----FRFHK 1536
 253 DSEVEFCFPREDPGAHRLMHIARLRLVLSPGWQLDAPCLPEKVTLCMQAPDPSQCP 312
 1537 EEKLMKICKIKNLSSSE-----IGQGAN----- 1560
 313 LVPPVPOKNATVNEPQDQVLVAGHPN--LCVOVSTWEKVOLOACSNADSLGPFKDMIL 369
 1561 ---VKAKMLT---KDRLSKKKKKWTGCIQVYIIT---VANCSDQJNPEGENSEI 1607
 370 VEMTGLNNTVSCALPEPGCTPLPSMASTRARALGEBLQ--DF--RSHQCM--QLAND 422
 1608 SEEYSKMWFTVSKSHGVEVYIHQ--KMKTEIGQCAIDISDYLKSHODLSKNSLWNS 1666
 423 -----DNMGSIMACPMKYIHRRWLVWLACILLAALEFFILKKDR 465
 1667 RSTVLNNSLSKSSIGVDKISLSKSPQEPHHYSAI-----ER 1705
 466 RKARSGSTALLSHSADAGYERLVGALASALSQMLPRAVAVDLSRRELSA----- 516
 1706 NMLRLSOTIPTFQVLPAGGEITYVRIEES--SPLMDKSMSSMSQGRAMIYVLSREE 1764
 517 -HGAL-----AMFHQRRIIDEGV--VILLESRAVAQCOQMLQLOTVERGPHDAL 566
 1765 MDGGLRRAMRVVSTWSEDD--ILKPGOVFIYKSFLEVVRTYHKKIFQESTV----- 1813
 567 AAMLSCVLPDFLOGATGATGYGVYFDGLLHPDSVSPFRVAPLE 610
 1814 ---LHLCLREIQOQRAAKL--ITYFNQVKKPOTIPTYPRLEVF 1852

RESULT 8
 55225 PRELIMINARY; PRT; 2910 AA.

01-JUN-1998 (TREMBLrel. 06, Created)
 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 OTOGELIN.
 Mus musculus (Mouse).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID-10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-BALB/C.
 RX MEDLINE-98070772; PubMed-9405633;
 RA Cohen-Salmon M., El-Amraoui A., Leibovici M., Petit C.;
 RT "Otogelin: A glycoprotein specific to the acellular membranes of the
 inner ear."
 RL Proc. Natl. Acad. Sci. U.S.A. 94:14450-14455(1997).
 DR EMBL: U96411; AAB96361.1;
 DR HSSP; P56682; ICV.
 DR MGD; MGI:1202064; Olog.
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR002919; TIL.
 DR InterPro: IPR001846; Wvd.
 DR Pfam; PF01826; TIL; 3.
 DR Pfam; PF00094; Wvd; 4.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00214; WVC; 1.
 DR SMART; SM00216; WVD; 4.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE; PS01225; CTCK_2; 1.
 SEQUENCE 2910 AA; 313410 MW; 7270FC61A23264CD CRC64;

Query Match 3.1%; Score 110; DB 11; Length 2910;
 Best Local Similarity 20.8%; Pred. No. 3.1;
 Matches 54; Conservative 31; Mismatches 82; Indels 92; Gaps 13;

QY 214 VPDALSKMYKMLTGQNTLTNTDVLPCICIQWLSLEPS-----ERVECCPR 263
 DB 2427 VPRALGETWNSLSG-----CCQOQCCAPDITIPVLDLDCPGRPRESCP-- 2469
 264 EDPGARMLMHIARLRLVLSPGWQLDAPCLPEKVTLCMQAPDPSQCPPLVPVQKNAT 323
 DB 2470 -----RGEVYILDP-----TEDPCL-GSYCV-----NOTLDEGLAPTCRPGHSL 2510
 QY 324 VNEPQDQVLVAGH-----PMLCVOVSTWEKVOLOACSNADSLGPFKDMILVEMKGLNN 378
 DB 2511 ITHQEDSCPSYSECDPGLC-----EAEQVPTC-----REDQILIEGLG-- 2552
 QY 379 TSVCALPEPGCTPLPSMASTRARALGEBLQ--DF--RSHQCM--QLAND 422
 DB 2553 DSCCTSYFCGGGECSDPMPECEGEGALTVHRTTELCCPLYQCVCFNFRCPQVOCGMCTS 2612
 QY 416 CMQMLNDNDMGSIMACPM 434
 DB 2613 LVYVWSPDRCPCPKSCED 2631

RESULT 9
 060150 PRELIMINARY; PRT; 361 AA.
 AC 060150;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)
 DE HYPOTHETICAL 40.1 KDA PROTEIN.
 GN SPBC18H10.20C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID-4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972H;
 RA Lyne M., Rajandream M.A., Barrell B.G., Badcock K., Churcher C.M.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF022304; CAA18417.1;

44 GDVLCLEPSIQ-----SAPGVLVPTRIQTELYLCRQKTDICALRVVVHLLA 91

QY 287 QLDAPCGLPKGVTLICWQAPDQSPQPLVPPVPPQKNATVNEPQDFQLVAGHPNLCVQYSTW 346


```

134 LARAPGOLPGTVPAHSHVLYAOEASALSAPAGDGLTLH-----LGSGD-----RLAHW 182
137 EKVOLOACSNADSLGPFDDMLVEMKTC--LNTSVCALEPGSGTLPSPMASTRARIG 404
183 QGGLL---SWR-SLSP--NMOLLVODETSQPPVGGQLALRPD-----LP--ASER----- 225
405 EELLQDRSHQCMLNDNDNGSLWACPMKDYTHRRVWLWLAACLLAALFFLLKKD 464
226 -QLFLTVGGQLOLVFSGD-----YLLRRR 250
465 -----RRKAARSTALLSHSADGAGYERLYGALSALSOPLRVAVLMSRRRLS 515
251 ADAAGTTHLRLALIGRACALLLPBQHGRLRLARTLARLNGDP-----PRADDP 302
516 AHGALAMFH-----HORRLIQEGGVILLFSPAAYVACCOMLQLOTFEBGPDALAAW 569
303 ASGOVAFAPGPAECLAARSLQLRLEPLGRSPAPVA-----HEAQVAQ 347
570 LSCVLPDLQGRATGRVGYFEDGLHPDSVSPFRAV---PLFSL---PTQLPA---FL 620
348 ALHLRP-----HTAAQVQAALHAASAPLEALATDPLVGRBAGPPLVATGHFI 399
621 DALQGCSTAGRPADRVRYTQALRSAL 649
400 VEPPLGPDPLEVRLEPGDRLLTLRPDYRAEL 428

```

RESULT 12

```

094G03 PRELIMINARY: PRT: 589 AA.
094G03:
01-DEC-2001 (TREMBLrel. 19, Created)
01-DEC-2001 (TREMBLrel. 19, Last sequence update)
01-DEC-2001 (TREMBLrel. 19, Last annotation update)
PUTATIVE CYTOSOLIC NADP-MALIC ENZYME (EC 1.1.1.40).
CYTMEA.
Flaveria pringlei.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; easterids II; Asterales; Asteraceae; Asteroideae;
Helentiae; Flaveria.
NCBI_TaxID=4226;
[1]
SEQUENCE FROM N.A.
Lal L.B., Nelson T.;
"Expression of the cytosolic NADP-malic enzyme genes in Flaveria.";
Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
EMBL: AF288920; FAK83073.1;
Oxidoreductase.
SEQUENCE 589 AA: 65124 MW: AB3AGFB01BBEP95A CRC64:

```

Query Match 3.0%; Score 106.5; DB 10; Length 589;
 Best Local Similarity 17.8%; Pred. No. 0.71;
 Matches 106; Conservative 70; Mismatches 182; Indels 237; Gaps 26;

```

135 LLEVOVPADLVNQGSGAVFQCFEASLGEVQI---WS-----YT---KRYQKEIN 182
136 LKQMGESGVVDDSSAVGGVDVYGEDRASEDLITPWTVSASGYTLRLDPHNHKGIA 64
183 LTVQQLPDGDNVLLTLVSEQDFSEFLYLRLVVPDALKSLMWKNLGNLPQNTLNHTLVPC 242
65 FTER-----ERDSHYLGLCLPPAVATQELQEKLM--QNRISTEYVPL--H 105
243 LCIQVWSLEPDSERVEFCPEDEPGAHNRNLHIALRYLSPGWOLDAPCCLPQKTYLWC 302
106 RYVAMMELEERERLEY-----KLLIDNVEEL-----LPVYVT--- 138
303 QAPDQSPQPLVPFPVQKNTV--NEPDQVLAVAGHPNLCVOV--STWEKYOLOACSNADSL 360
139 -----PTVGEACQKYGSIKRPQGLYISLKEKGLILEVLRNMPFNINQIVYVTDG- 188

```

```

QY 361 GPFKDDMLVEMKTCGLNN-----TSVCALEPGSGCTLPSPMASTRARL- 403
DB 189 -----ERLIGLGDGCGMGIPVGLALVATLGGVPSACLPITDIVGTNNOKLL 238
QY 404 -----GEELL---QDRSHQCML----- 419
DB 239 DDEFTIGLRKATGKEYDYDLLEEFMSAYKQNGEYKLVQPEDANNAFLAKYRTSH 298
QY 420 --WDDNNGSLWACPMKDYTHRRVWLWLAACLLAALFFELLKKRRKARCSRTALL 477
DB 299 LVFNDDIGT-----ASVYLAGVLSKLL-----GGSIDHTLFL 334
QY 478 LHSAD--GAGYERLYGALSALSOPLRVA--VDIM-----SRRLSAHGALAMFH 525
DB 335 LGAGEAGTGAIELIALEISTKANIPIDEARKITWLDKGLVSSRKETLOHFKPMAHE 394
QY 526 -----ORRLIQEGGVILLFSPAAYVAC-- 549
DB 395 HEPLSTLIDAKRAIPSVLIGTSGVQOTFTQNVVEAANERPLIMALNPTSOAECTA 454
QY 550 ---QWLOLOQTV-----EPGPDALAAWLSCLVLPDLQGRATGRVGYFEDGL 594
DB 455 EQAVTWTQGRAIFFSSGSPDEPYEYNGNL-----FIRGQANNAYI---FPGL 497

```

RESULT 13

```

Q9UJ61 PRELIMINARY: PRT: 1248 AA.
Q9UJ61:
01-MAY-2000 (TREMBLrel. 13, Created)
01-MAY-2000 (TREMBLrel. 13, Last sequence update)
01-DEC-2001 (TREMBLrel. 19, Last annotation update)
APOPTOTIC PROTEASE ACTIVATING FACTOR 1.
APAF-1.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
Hahn C., Duerkop H., Stein H.;
"Sequence of Apaf-1 in L.428.";
Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
EMBL: AJ243009; CAB55585.1;
HSSP: 014727; ICY5.
InterPro: IPR001315; CARD.
InterPro: IPR000767; Disease_resist.
InterPro: IPR002182; NB-ARC.
InterPro: IPR001680; WD40.
PFam: PF00619; CARD; 1.
PFam: PF00931; NB-ARC; 1.
PRINTS: PR00364; DISEASESERSIST.
SMART: SM00320; WD40; 11.
PROSITE: PS50209; CARD; 1.
PROSITE: PS00678; WD_REPEATS_1; UNKNOWN.
PROSITE: PS50082; WD_REPEATS_2; 9.
PROSITE: PS50294; WD_REPEATS_REGION; 1.
Protease; Repeat; WD repeat.
SEQUENCE 1248 AA: 141728 MW: DDA7285E6D5C8F37 CRC64:

```

Query Match 3.0%; Score 106.5; DB 4; Length 1248;
 Best Local Similarity 19.1%; Pred. No. 2;
 Matches 112; Conservative 64; Mismatches 188; Indels 221; Gaps 27;

```

QY 79 DCALRVVVVHLAVHGH-----WAPEEAGKSDSELOE--SRNASLOA--QVLSFOY-- 128
DB 537 DCAVSENFQELSLNGHLGNQFPFNIVQIGCEPETSIVYQQAQAOEYVNGHLYLE 596
QY 129 -----PLARCALLLEVQVPADLV-----QPGSGVGSANFV---CFEASIG----- 164

```


597 WINKKNTNLSCLVVRPHPTDVAHYHACFSEDEGRIASCGADKTLQVFAETGEKLEIKAH 656
165 -----AEVOIMS-----YTKPROKELNLOOLPDGDVNL 195
657 EDEVLCACFSTDOFIATCSVDKAKVIMNSMTGELVHYTHDESEOVNCHFTNSSHLL 716
196 TLDSVEQDSFLLYLVRPVDALSKMYKNLGPONITLNLTLVPCLCIQVMSLEPDS 255
717 ATGSS-----DCFLKMLDNOKECNTMFGHTN----- 744
256 RVEFCFPREDPGAHNIMHARLRYLSPGWOLDAPCCLPKVTLCMOADPSCOPLPV 315
745 SVNHCFRSPD-----KILASCSADGTLKL-WDATSANERKSI-- 781
316 PVPQKATVNEPDQFOLVAGHPNLGVSTWKKYQLOACSW-ADS-----LGP 362
782 NVKQFPLNEDPDQDM-----EVLYKCCSWSADGARINVAANKKIFLEA 825
363 FKDDMLLVKKTGLNNT-SVCALEPSCGTPPLPSMASTRARALGELLQDRSHOCMOAMN 421
826 IHTSGLLGELIHTGHSTIOYCDFSPQNLAVVNL-----SOTCVELMN 868
422 DDMGSLMACPMKYIHRMYLVMLACLLLAALFEFLILKDRRRKARSRTALLHSA 481
869 TDSRSKVDRCG-----HLSMV-----HG-----VWFSP 892
482 DQGYERLVNALSALQMPRAVADLMSRRELASGALAMPHHRRRIIOEGVILLF 541
893 DGSSF-----LTSSDDQ-----TIRLETKKYCKNSAV-MLKQEVVYQENEMVVL- 938
542 SPAAVACCOOQLOLQVPEPHDAL--AAMLSCVLPDLOGRATG 584
939 ---AVDHIR-LQLVNGRTGOIDYLTGAOVSCCCLSPHLQYIAFG 979
ESULT 14
93XR9
093XR9 PRELIMINARY; PRT: 756 AA.
01-DEC-2001 (Tremblrel. 19, Created)
01-DEC-2001 (Tremblrel. 19, last sequence update)
01-DEC-2001 (Tremblrel. 19, last annotation update)
BG55 PROTEIN.
Bruguiera gymnorhiza.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids I; Malpighiales; Rhizophoraceae; Bruguiera.
NCBI_TaxID=39984;
[1]
SEQUENCE FROM N.A.
TISSUE=LEAF;
Banazai T., Hershkovits G., Katcoff D.J., Hanagata N., Dubinsky Z.,
Karube I.;
Identification of mRNA transcripts differentially expressed in
response to high salinity by means of differential display in the
mangrove plant, Bruguiera gymnorhiza.
Submitted (May-2001) to the EMBL/Genbank/DBJ databases.
EMBL: AB061795; BAB55653.1;
SEQUENCE 756 AA; 84482 MW; 1597F3D46C3E0809 CRC64;

Db 84 NGDIVSRCKQSRNMLEOSLDQIQTMVPVILAAEISQVIDDLVANKVILDH--SDEEAGK 141
QY 105 SDSELOE---SRNASLOAOV-VLSQAVPIARCALLLEVOPADLVNPGQSVGSAVPCFE 160
Db 142 AVRELLDOHTMSDSDVESIEIKVLRPA-----SRLCITTPKOLLIERKSI----- 186
QY 161 ASLGAEOIWSYTKPRYKELNLOOLPDGDVNLTLDSVEQDSFLLYLVRPVDALKS 220
Db 187 -----KLVKVKR-----DN-----DPKKKILYLLH-----LKK 213
QY 221 LMYKNLGPONITLNH-----TDLVPC-LCIGV-----MSLEPDSERVE 258
Db 214 -YGNSTLGEQENLNSQOELFPADGSSVSQAAEVGPCTACKQIYAAEKNISNIPAPPE 272
QY 259 F-CFPR-----DPCAHNIMHARLRYLSPG-----VWOLDAPCCLPKVTLCMOADPQS 308
Db 273 YKCPLSKRLMYP-----VVIASGQFERIM-----IQKWFDEGND 308
QY 309 PC-----QPLVPVPQKATVNEPDQFOLVAGHPNLGVSTWKKYQLOACSWADS 359
Db 309 TCPEKTLVKLDHOSIMPTALKLISKWCEKYGVTIIDPN---SQAFOSLDTSSTVA-S 363
QY 360 LGPEKDDMLT---VEMKTGLNNTSVCALEP-----SGCTPLPSMAST- 398
Db 364 FGISMNDLHLPLDISNVLSGSDASYSQSPRTKIAERSMIMSMORNGYSAPFSRANTN 423
QY 399 -----RAARLGE---ELQDRSHOCQLOMNDNGSLMACPMKYIHRMYLVML 446
Db 424 KTCIDLPFSRLAKLGESKCEMEDVKSILH-----EDNVQPFHHSIENERVE 469
QY 447 ACILLAALFEFLILK---DPRKARSGRTALLHLSADAG-----YERLVGALAS 495
Db 470 -----PLIKFLDKAQOHVRAORAGSKLLAEVSKRSGISMLHEDPDLIASMLDS 522
QY 496 ALSQMPLRV---AVDLSRRELASGALAMPHH---QRRRIIOEGVILLFSPAQAQ 548
Db 523 ELVEBALAILLEVLSDDKRSKRITASGALVYIIRIDSEREEQEGAVRIILHNSNNEV 582
QY 549 COQWLOLQVPEPHDALAMLSCVLPDLOGRATGRYGV 589
Db 583 CSQILSINCIPK-----LVFPINQOLASHCMQL 611
RESULT 15
ID 09R6L2 PRELIMINARY; PRT: 1693 AA.
AC 09R6L2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE TIORE47 PROTEIN.
GN TIORE47.
OS Agrobacterium tumefaciens.
OC Plasmid pTI-SAKURA.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=358;
ON [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RX MEDLINE=20184752; PubMed=10721727;
RA Suzuki K., Hattori Y., Uraji M., Ohta N., Iwata K., Murata K.,
Katoh A., Yoshida K.;
RT "Complete nucleotide sequence of a plant tumor-inducing Ti plasmid."
RL Gene 242:331-336(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RX MEDLINE=98193120; PubMed=9524202;
RA Suzuki K., Ohta N., Hattori Y., Uraji M., Katoh A., Yoshida K.;
RT "Novel structural difference between nopaline- and octopline- type trbJ
gene: construction of genetic and physical map and sequencing of
trbJ/tral and rep gene clusters of a new Ti plasmid pTI-SAKURA.";

BL0chlm. Biophys. Acta 1396:1-7(1998).

[13] SEQUENCE FROM N.A.
 RC STRAIN-MAFE301001;
 RC Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;
 RT "Genome structure of pTI-SAKURA(I): Strategy for DNA sequencing of a
 RT Japanese cherry-ti plasmid.";
 RT Nucleic Acids Symp. Ser. 37:159-160(1998).
 [14] SEQUENCE FROM N.A.
 RC STRAIN-MAFE301001;
 RC Ohta N., Suzuki K., Hattori Y., Uraji M., Katoh A., Yoshida K.;
 RT "Genome structure of pTI-SAKURA (III): Characteristics of T-DNA.";
 RT Nucleic Acids Symp. Ser. 39:185-186(1998).
 [15] SEQUENCE FROM N.A.
 RC STRAIN-MAFE301001;
 RC Uraji M., Suzuki K., Ohta N., Hattori Y., Katoh A., Yoshida K.;
 RT "Genome structure of pTI-SAKURA (IV): Characteristics of tra region.";
 RT Nucleic Acids Symp. Ser. 39:187-188(1998).
 [16] SEQUENCE FROM N.A.
 RC STRAIN-MAFE301001;
 RC Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;
 RT "Genome structure of pTI-SAKURA (V): Complete nucleotide sequence of
 RT plasmid pTI-SAKURA's vir region in Agrobacterium tumefaciens.";
 RT Nucleic Acids Symp. Ser. 39:265-266(1998).
 RC EMBL: AB016260; BAA87672.1;
 RC HSSP: P14385; 2ADM.
 RC InterPro: IPR001650; Helicase_C.
 RC InterPro: IPR002296; N12N6_mltifase.
 RC InterPro: IPR002032; N6_Mtase.
 RC InterPro: IPR000051; SAM_bind.
 RC Pfam: PF00271; Helicase_C; 1.
 RC PRINTS: PR00507; N12N6MTFRASE.
 RC PROSITE: PS00092; N6_MTASE; UNKNOWN_1.
 RC ATP-binding: Helicase; Plasmid.
 RC SEQUENCE 1693 AA; 187978 MW; 0DD679AFBE6514BE CRC64;

Query Match 2.9%; Score 106; DB 2; Length 1693;

Best local Similarity 19.4%; Pred. No. 3.4; Mismatches 159; Indels 254; Gaps 32;

Matches 115; Conservative 66; Mismatches 159; Indels 254; Gaps 32;

79 DCALRVVVLAVHGM-AEPEEAGKSDSELOESRNASLQAOVLSFOAYPIARCALLE 137
 414 DLVVALSAVHLPEGHGDETEI-DLDELEATDA-----P 450
 138 VOVPADLVPGQSV--GSAVDFCEFAEVOIWSYTRPYOKELNLTQOLPPGDVYL 195
 451 VDLPSD-----RHVREGSFFPDNARGLM-----QVIDGEPAVL 483
 196 TL-----DYSEEDFSFLKLRPYDALKSLM-----YKNLT----- 227
 484 KYKGRNADGILEHVRIRLIPRDVREVLAQELDRPMKDLQRLRVAMSSFYRDF 543
 228 GPONITLNFHDLVPCLCIOVNSLEPDS-----RYEFCFEREDGARNLMHIALRL 278
 544 GP-----INHITV-----SITEDPESGETRESHRPNLOPFADDP---DCMLVASI 586
 279 RVLSPGVQWLDAQCCLPQKVTLCMQAPDQSPCOPLVPYPQKNATVNEPQDFQVAGHPN 338
 587 E-----DYDLENDTAKPGAIF-----TERVISPPAPV-----ITSADA 621
 339 LCVQVSTWEKVOQAC-----SW--ADSL--GPFKDDML 368
 622 LAVVLENGRVVDLDHIAELLHRDDVVAELGSAIFRDPADGSMQADAVLSGVRDLK 681
 369 LVEMKTGLN---NTSVCALPESGCTPLPSPMASTRARLG-----ELLQDF 411
 682 VAEEAALADPYNRNRYAL--AGVQPYDLRPSDITARLGAPWIPADVVAFVKEMGTDI 739
 412 RSHQCMQ--WDD--NMGSIMACPMKVIHRRNVLVWLACLLAALFFELLKKDRRK 467

Db 740 RIHHMPELAWVTEARQGLACTSE-----W-----GTDRRH 773
 QY 468 A-----ANGSRTALLHS-ADGAGYERLVGALASALQMPLRVAVD-----LWS----- 510
 Db 774 AGELSDALNSRVPQJFDFTIRGDSERRVLNVVDTEAAKEKELKIKDAFQRMWSDPDR 833
 QY 511 -----RRELS-----AGCALMFHHQR--RILOEGGAVI 538
 Db 834 DLARVYNDRENNIAPRKESGDHLNLPAGSAGFVLYGHQKRGIMRISSGTYL 887

Search completed: September 28, 2002, 19:42:54
 Job time: 1007 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

protein - protein search, using sw model

on: September 28, 2002, 19:38:52 ; Search time 24.51 Seconds
(without alignments)
1064.748 Million cell updates/sec

US-09-899-471-2
ect score: 3605
nce: 1 MPVSWEFLSLALGRNPVVS.....SSEAPGCCCEMDLPCYTLE 674

ing table: BIOSUM62
Gapop 10.0 , Gapext 0.5

ched: 105224 seqs, 38719550 residues

1 number of hits satisfying chosen parameters: 105224

um DB seq length: 0

um DB seq length: 200000000

rocessing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ase : SwissProt_40:*

SUMMARIES

Score	Query	Match	Length	ID	Description
116.5	3.2	1711	1	PPPO_RAT	O6612 ratu
105.5	2.9	1451	1	A2M2_MOUSE	P28666 mus musculu
105	2.9	1443	1	W70T_RAT	O35828 ratu
104.5	2.9	1248	1	APAF_HUMAN	O14727 homo sapien
103.5	2.9	3707	1	PGM_MOUSE	O05793 mus musculu
100.5	2.8	1062	1	NAL2_HUMAN	O9nx02 homo sapien
100	2.8	1447	1	DCC_HUMAN	P43146 homo sapien
99.5	2.8	2647	1	ABR2_HUMAN	P21333 homo sapien
99	2.7	922	1	W70T_MOUSE	Q9d2v7 mus musculu
99	2.7	2437	1	NOTC_BRARE	P46530 brachydantio
98.5	2.7	1447	1	DCC_MOUSE	P70211 mus musculu
97.5	2.7	1618	1	NEST_HUMAN	P48681 homo sapien
97	2.7	963	1	URB4_HUMAN	Q13107 homo sapien
96.5	2.7	1476	1	RVR3_HUMAN	Q15413 homo sapien
94.5	2.6	837	1	GCGR_MOUSE	P28665 mus musculu
94.5	2.6	1587	1	LMG3_HUMAN	P40223 mus musculu
94	2.6	962	1	URB4_MOUSE	Q9y6n6 homo sapien
93.5	2.6	204	1	C267_HUMAN	P35123 mus musculu
93.5	2.6	1914	1	STCK_EMENT	P88512 homo sapien
93.5	2.6	2390	1	SPCP_HUMAN	O00706 emeticella
92.5	2.6	1249	1	APAF_RAT	O15020 homo sapien
92	2.6	398	1	GPPL_KLEPN	Q9epv5 rattus norv
92	2.6	1912	1	PRPD_HUMAN	P15751 klebsiella
91.5	2.5	235	1	FLJ31_HUMAN	P23468 homo sapien
91.5	2.5	461	1	NIFN_KLEPN	P49721 homo sapien
91.5	2.5	486	1	ABR3_HUMAN	P08738 klebsiella
91.5	2.5	880	1	TYO3_MOUSE	O55704 homo sapien
91.5	2.5	890	1	TYO3_HUMAN	P55144 mus musculu
91.5	2.5	1073	1	VAS5_SCHPO	O06418 homo sapien
91	2.5	1208	1	RCO4_HUMAN	O10141 schizosacch
91	2.5	362	1	CKRA_HUMAN	O94761 homo sapien
91	2.5	505	1	CXAA_MOUSE	P46092 homo sapien
91	2.5	505	1	CXAA_MOUSE	O9wus4 mus musculu

34	91	2.5	1104	1	NFX1_HUMAN	O12986 homo sapien
35	91	2.5	1189	1	HAIR_HUMAN	O43593 homo sapien
36	91	2.5	3695	1	IMA5_HUMAN	O15230 homo sapien
37	91	2.5	3065	1	EPPL_HUMAN	P58107 homo sapien
38	91	2.5	385	1	YEHY_ECOLI	P33361 escherichia
39	90.5	2.5	458	1	MOC_RABIT	P03988 oryctolagus
40	90.5	2.5	479	1	MUCM_RABIT	P04321 oryctolagus
41	90.5	2.5	690	1	KATB_SCHRO	O10156 schizosacch
42	90.5	2.5	880	1	TYO3_RAT	P55146 rattus norv
43	90	2.5	537	1	ICAL_MOUSE	P13597 mus musculu
44	90	2.5	796	1	SG2N_MOUSE	O9erg2 mus musculu
45	90	2.5	1928	1	LPN_RAT	O02401 rattus norv

ALIGNMENTS

RESULT	1	STANDARD	PRT	1711 AA
PPPO_RAT				
ID	PPPO_RAT			
AC	O6612			
DT	01-NOV-1997	(rel. 35, Created)		
DT	01-NOV-1997	(rel. 35, last sequence update)		
DT	15-JUL-1999	(rel. 38, last annotation update)		
DE	Osteosticellular protein tyrosine phosphatase precursor (EC 3.1.3.48)			
DE	(OSF-PPP)			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
OX	NCBI_Taxid=101116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-Osteosarcoma;			
RX	MEDLINE=95074080; PubMed=7527035;			
RA	Mauro L.J., Olmsted E.A., Skrobacz B.M., Mourey R.J., Davis A.R.,			
RA	Dixon J.E.;			
RT	"Identification of a hormonally regulated protein tyrosine			
RT	phosphatase associated with bone and testicular differentiation.";			
RL	J. Biol. Chem. 269:30659-30667(1994).			
CC	-1- FUNCTION: MAY FUNCTION IN SIGNALING PATHWAYS DURING BONE			
CC	REMODELING, AS WELL AS SERVE A BROADER ROLE IN CELL INTERACTIONS			
CC	ASSOCIATED WITH DIFFERENTIATION IN BONE AND TESTIS. OPTIMAL PH FOR			
CC	PHOSPHATASE ACTIVITY IS 5.6.			
CC	-1- ASSOCIATED WITH DIFFERENTIATION IN BONE AND TESTIS.			
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O -> protein			
CC	tyrosine + phosphate.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-1- ALTERNATIVE PRODUCTS: A PRESUMED ALTERNATE TRANSCRIPT OF 4.8-5.0			
CC	KILOBASES, WHICH MAY LACK PPP DOMAINS, IS PRESENT IN PROLIFERATING			
CC	OSTEOBLASTS, BUT NOT DETECTABLE AT OTHER STAGES.			
CC	-1- TISSUE SPECIFICITY: BONE AND TESTIS. IN THE LATTER, RESTRICTED TO			
CC	THE BASAL PORTION OF THE SEMINIFEROUS TUBULE.			
CC	-1- DEVELOPMENTAL STAGE: UP-REGULATED IN DIFFERENTIATING CULTURES OF			
CC	PRIMARY OSTEOBLASTS AND DOWN-REGULATED IN LATE STAGE MINERALIZING			
CC	CULTURES. IN TESTIS, EXPRESSION IS HIGHEST BETWEEN STAGES I AND			
CC	VII WHEN MATURING SPERMATIDS REMAIN BURIED WITHIN THE SEPTOLI			
CC	EPITHELIUM.			
CC	-1- INDUCTION: BY PARATHYROID HORMONE AND CYCLIC AMP ANALOGS.			
CC	-1- PTM: THE CYTOPLASMIC DOMAIN CONTAINS POTENTIAL PHOSPHORYLATION			
CC	SITES.			
CC	-1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.			
CC	-1- SIMILARITY: CONTAINS 10 FIBRONECTIN TYPE III-LIKE DOMAINS.			
CC	-1- THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sdb.ch/announce/			
CC	or send an email to license@sdb-sdb.ch).			
CC	EMBL; L36884; AAA63911.1; -			
DR	HSSP; P18052; 1YFO.			

InterPro: IPR000340; DS-phosphatase.
InterPro: IPR003961; FN-III.
InterPro: IPR003962; FN-III-repeat.
InterPro: IPR000387; TYR-phosphatase.
InterPro: IPR000242; TYR-prot-phosphatase.
Pfam: PF00041; fn3.7.
Pfam: PF00102; Y-phosphatase; 1.
PRINTS: PR00104; PTRYPHIT.
PRINTS: PR00700; PTRYPHITASE.
SMART: SM00060; FN3.6.
SMART: SM00194; PTRYC.1.
SMART: SM00102; PTRYC-DSPC.1.
PROSITE: PS00383; TYR-PHOSPHATASE_1; 1.
PROSITE: PS00056; TYR-PHOSPHATASE_2; 1.
PROSITE: PS50055; TYR-PHOSPHATASE_PTP.2.
Hydrolase; Transmembrane; Repeat; Signal; Glycoprotein.
SIGNAL 1 17
CHAIN 18 1711
DOMAIN 18 1074
TRANSMEM 1075 1095
DOMAIN 1096 1711
CYTOPLASMIC (POTENTIAL).
DOMAIN 32 124
DOMAIN 125 215
DOMAIN 216 303
DOMAIN 304 392
DOMAIN 393 470
DOMAIN 471 562
DOMAIN 563 652
DOMAIN 653 741
DOMAIN 742 830
DOMAIN 831 921
DOMAIN 1150 1418
DOMAIN 1419 1711
ACT_SITE 1350 1350
CARBOHYD 42 74
CARBOHYD 74 89
CARBOHYD 89 117
CARBOHYD 117 174
CARBOHYD 174 239
CARBOHYD 239 259
CARBOHYD 259 431
CARBOHYD 431 431
CARBOHYD 570 570
CARBOHYD 620 620
CARBOHYD 649 649
CARBOHYD 663 663
CARBOHYD 737 737
CARBOHYD 851 851
CARBOHYD 882 882
CARBOHYD 970 970
CARBOHYD 982 982
SEQUENCE 1711 AA: 187292 MW: 0F04D2D1A47A18A0 CRC64;
Query Match 3.28; Score 116.5; DB 1; Length 1711;
Best Local Similarity 20.74; Pred. No. 0.39; Indels 267; Gaps 40;
Matches 150; Conservative 80; Mismatch 227;
43 DGDVLCPLSGLSA-----PGVLPVPTRIQTEL-VLRCPOK-----TDC 80
333 DGVYKLSGMESTSLGPECNAVFPGP-LPFGHYTLQKLAKYDAMWEGSTWLAES 391
81 ALRVRYVHLAVGHNAPEPEAGSDSELQSRNALSQAVLSFOAYPTARCALLLEVQV 140
332 AALPREVPGARL---WLDGLEASK---QGRRLALYSD-----DAPSLGINSV 434
141 PADLYO-----PGQ-----SVGSVDFCCFEASIGAEVQIMSYTKPRQKELNLQGLPD 189
435 PSGATVIVICGLVPGAHRYNDIASSTGDISISG-----YISDPLPOSLEV----- 481
190 GNVVLTLLDVSEODESFILYLRVPDAL---KSLMYKNLT-----GQONTLNH 236

Db 482 -----ISRSPSDLTIAMGPAPQLEGKYVTHHQDSQRSPGDLVDLGPDLSLTL 532
QY 237 TDLYP-CLCLOVW-----SLEPDSERVEFCP-----FREDGARNLWH----- 274
Db 533 KSLVPGSCYTVSAMAMAGNIDSQKIHSCTRAPPTNLSTGFAGHPALAKSWHPGCG 592
QY 275 -----IARLVL---SPGV-----W-OLDAPCCLPKGYTLCKQADQSCQPLV 314
Db 593 RDAFLRLRLRLPLLESKVLPREAONFSMAQLTACCEFOVQLSTLWGSRRS----- 646
QY 315 PPVPOKNATVNEPODFQVLAHGNLCVQVSTWEKVOQAQSMADSLGPFKMDMLYEMKT 374
Db 647 -----SANATGTP-----PSAPTL-VVNTSDAPTLQV-SWAVPG----- 681
QY 375 GLNNTSYCALPESGCTPLPSMASTRARLGEELLQDFRSHQCMQMDNMGSLWACPM 434
Db 682 GRSRQVLYLQESTRTATSTMP-----KEDGSFGLGTGT 718
QY 435 KYHRRVYLVNACILLAA---LFFELLKRRKRRKARSRTALLHSADAGTERLV- 490
Db 719 KY-KVEYISWAGPLYTAANVSAMTYPLIPNE-----LTVSMQGSVAVNLAM 765
QY 491 -----GALASALS-----OMPLRVAVDLSRRELASHGALAMFHORRRLIQEGV 536
Db 766 PSGFLGQCAHQSLSDAGHLSWEQPLKQELMLNDLTPGHYTSVSCRAGPLQASTH 825
QY 537 VILFSPAAVACQOMLOLQVPEGP-HDAL-----AAMLS-----CVLPDLQ 579
Db 826 LVVL-----SVEPGVEDYLCHPEFTYLLNMTMPAGDVVCLV----- 864
QY 580 GRATGRVY---GVYF-----DGLHPDSVP-SPPRVAPLSPLQPLPAFLDALQGC 627
Db 865 --VVERLEVPGGGTHRVFOVNTSGDALILPLNMTPTSYRLS--LTVLCRNSWMSNVLVC 920
QY 628 STSA 631
Db 921 STSA 924
RESULT 2
ID A2M2.MOUSE STANDARD; PRT: 1451 AA.
AC P28666;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Murioglobulin 2 precursor (MUG2).
GN MUG2 OR MUG-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Liver;
RX MEDLINE=91358495; Pubmed=1840592;
RA Overbergh L., Torrekens S., van Leuven F., Van den Berghe H.;
RT "Molecular characterization of the murine Murioglobulin 2";
RJ J. Biol. Chem. 266:16903-16910(1991).
CC -FUNCTION: A PROTEINASE ACTIVATES THE INHIBITOR BY SPECIFIC
CC PROTEOLYSIS IN THE BAIT REGION, WHICH, BY AN UNKNOWN MECHANISM
CC LEADS TO REACTION AT THE CYSTEINYL-GLUTAMYL INTERNAL THIOLESTER
CC SITE AND TO A CONFORMATIONAL CHANGE, WHEREBY THE PROTEINASE IS
CC TRAPPED AND/OR COVALENTLY BOUND TO THE INHIBITOR. WHILE IN THE
CC TETRAMERIC PROTEINASE INHIBITORS STERIC INHIBITION IS SUFFICIENTLY
CC STRONG, MONOMERIC FORMS NEED A COVALENT LINKAGE BETWEEN THE
CC ACTIVATED GLUTAMYL RESIDUE OF THE ORIGINAL THIOLESTER AND A
CC TERMINAL AMINO GROUP OF A LYSINE OR ANOTHER NUCLEOPHILIC GROUP ON
CC THE PROTEINASE, FOR INHIBITION TO BE EFFECTIVE.
CC -SUBUNIT: MONOMER.
CC -TISSUE SPECIFICITY: PLASMA.
CC -SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
CC INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: M65238; AAA73041.1; -
PIR: B41185; B41185.
HSSP: P01023; 1BVB.
MGP: MGI:99836; M92.
InterPro: IPR002890; Alpha_2_macroglabin.
Pfam: PF00207; A2M; 2.
Pfam: PF01835; A2M; N; 1.
ProSite: PS00477; ALPHA_2_MACROGLABIN; 1.
Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;
Multigene family.
SIGNAL 1 27 BY SIMILARITY.
CHAIN 1 28 1451 MURINOGLABIN 2.
DOMAIN 677 734 BAIT REGION (APPROXIMATELY).
DISULFID 48 86 BY SIMILARITY.
DISULFID 251 276 BY SIMILARITY.
DISULFID 269 288 BY SIMILARITY.
DISULFID 461 555 BY SIMILARITY.
DISULFID 587 773 BY SIMILARITY.
DISULFID 634 680 BY SIMILARITY.
DISULFID 849 885 BY SIMILARITY.
DISULFID 923 1274 BY SIMILARITY.
DISULFID 1081 1104 BY SIMILARITY.
DISULFID 1298 1444 BY SIMILARITY.
DISULFID 974 977 BY SIMILARITY.
DISULFID 55 55 N-LINKED (GLCNAC...) (POTENTIAL).
DISULFID 294 294 N-LINKED (GLCNAC...) (POTENTIAL).
DISULFID 313 313 N-LINKED (GLCNAC...) (POTENTIAL).
DISULFID 500 500 N-LINKED (GLCNAC...) (POTENTIAL).
DISULFID 749 749 N-LINKED (GLCNAC...) (POTENTIAL).
DISULFID 776 776 N-LINKED (GLCNAC...) (POTENTIAL).
DISULFID 871 871 N-LINKED (GLCNAC...) (POTENTIAL).
DISULFID 1401 1401 N-LINKED (GLCNAC...) (POTENTIAL).
SEQUENCE 1451 AA; 162368 MW; 982A82EFAA6258F6 CRC64;

Query Match 2.9%; Score 105.5; DB 1; Length 1451;
Best Local Similarity 20.4%; Pred. No. 2.5;
Matches 76; Conservative 48; Mismatches 159; Indels 89; Gaps 17;
67 QTELVLCRQ-----KTDCALEVRVYHLAV-----HGWAPEEA--- 102
454 ETSSILPCNQITVQAHFLIKGDLGVLKELVYFYLVMAQSIIGTGNTHOYEGEAPRK 513
103 GKSPSELOESRNASQAOVYVLSQAPPIARCALLEVVPADLVPGOSVGSVAVDFEAS 162
514 GNFDLE-----IPVFSNAPMAKMLYTTLPDGEVI--ADSVNFEIEKCLRNRK 559
163 LGAEVQISYTKPRYKELNLTQQLPDGDNVLTLDVSEBDSFLLYLKRPVPAKSLM 222
560 VDLSSSS--SQSLPACSTRQAVTAS--PQSLCGLNAVDS-----VLLKRP-EDSLSPSM 609
223 YKNTLGPONTLNTHTDLV-----CLCLOVMSLEPDSERVEPCPFREDPAHR-- 270
610 IYNLPGMQ-----HNKFISSLSLEDRDCILYSSWAE--KHTDWPGRREKRYRIV 661
271 ---NLNHLARLAVLSPGWQDLAPCLP-GKVTLCWQAPDQSPQ-----PLVPPVOK 320
662 EDMCLKAETNLIKIKIPKICFDSAPMGPRGKFDLAFSSVSGTLQKSSSKRPQREPPRE 721
321 NATVNEPQDFOLVACHPNCVQVSTWEKVOLOACSN-----DSLCPFFDDMLVEMK 373
722 DPEKDPDLAETIRKTFPEETW-----MDIVTVNSTGVAEVEVTVPDTITTEKAGALCLSD 777

OY 374 TGLNNTSCALE 385
DB 778 TGLGLSSVPLQ 789

RESULT 3
W70T_RAT STANDARD; PRT; 443 AA.
ID W70T_RAT
AC O35828;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE 70 kDa WD-repeat tumor rejection antigen (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR;
RX PubMed-9703019;
RA Muralikrishna T., Begum Z., Swamy C.V., Khar A.;
RT "Molecular cloning and characterization of a tumor rejection antigen
from rat histiocytoma, AK-5".
RL DNA Cell Biol. 17:603-612(1998).
CC -1- SIMILARITY: BELONGS TO THE CORONIN FAMILY OF WD-REPEAT PROTEINS.
CC -1- SIMILARITY: CONTAINS AT LEAST 4 WD REPEATS (TRP-ASP DOMAINS).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

Query Match 2.9%; Score 105; DB 1; Length 443;
Best Local Similarity 20.0%; Pred. No. 0.56;
Matches 91; Conservative 50; Mismatches 137; Indels 176; Gaps 21;
OY 60 VLVPRLTQTELVLCRQKTD--CALRVVVVYHLAVHGWAPEEAGKSDSELOESRNAST 117
DB 2 IRVPCNTLKGMLNTTGESDGFPCANRLRVAVPLLSG----- 38
OY 118 QAQVLSFGAYPIARCALLEVVPADL-----VQPCDSVGSVAVDFE-----ASLG 164
DB 39 -----GOVAVLELQKPRGLPDTGLPTLQNGAAVMDLVWDFDPRLAVAGED 85
OY 165 AEVQISYTKPRYKELNLTQQLPDGDNVLTLDVSEBDSFLLYLKRPVPAKSLM 224
DB 86 ARIRLM----- 101
OY 225 NLGPGONTLNTHTDLV-----PC-----LCLOVMSLEPDSERVEPCPFREDPG 267
DB 102 -LTPEETVLTGHTKIKYSLRHLADVLASSYDLTIRINDLQARTQRL-----QG 154
OY 268 AHRNHLARLAVLSPGWQDLAPCLP-GKVTLCWQAPDQSPQPLVPPVPO-KNATVNE 326


```

b 155 HODQIETSLA-----WSPDGKQL-ATYCKDKRVR--YDRPSSPLPLQDGEPEBEGRGAASS 207
      : : : : : | | | | | : : : : :
y 327 PODEFOLVACHPNLCVOVST-----NEKYOLOACSNADSLGP-----FKDDMLIVE 371
      : : : : : | | | | | : : : : :
b 208 GCMVYVVCYUUVILTAEVASCSCTMLMPQOKA-LQHC-WASTYAPSTILSPYDPDGLV- 264
      : : : : : | | | | | : : : : :
y 372 MKTGILNNTSVSCALE-----PSCCTPLPSMASTRARLGE 405
      : : : : : | | | | | : : : : :
b 265 LITGKGDIVFLYEIVIRPEARFLECSFTSPDRHKGFILLPKTECIDIQVEFARCLRLQ 324
      : : : : : | | | | | : : : : :
y 406 ELLQD--FRSHOCMOLMND---NMGSLWACRMD 434
      : : : : : | | | | | : : : : :
b 325 TSLEPVAERLPRVREKFEQDDVPEDTAVTASALMD 358
      : : : : : | | | | | : : : : :

```

RESULT 4	PAF HUMAN
----------	-----------

Cell 90:405-413(1997).

SEQUENCE FROM N.A. (ISOFORM 2).
TISSUE-T-cell;
MEDLINE-99292765; Pubmed-10366241;
Saleh A., Strinivasula S.M., Acharya S., Fishel R., Alnemri E.S.,
Cytochrome c and dAMP-mediated oligomerization of Apaf-1 is a
prerequisite for procaspase-9 activation.,"
J. Biol. Chem. 274:17941-17945(1999).

SEQUENCE FROM N.A. (ISOFORM 1).
TISUE=Kidney;
Benedict M.A., Nunez G.;
Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases

TISSUE-BRAIN;
MEDLINE-98116555; PubMed-9455477;
Ishikawa K.-I., Nagase T., Nakajima D., Seki N., Ohira M.,
Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
Prediction of the coding sequences of unidentified human genes. VIII.
78 new cDNA clones from brain which code for large proteins in
vitro. *J. Mol. Biol.* 1997; 270:123-34.
DNA Res. 4:307-313(1997).

[6] SEQUENCE OF 810-864 AND 866-883 FROM N.A. Roberts D.L., Daigleish R., Cohen G.M., MacFarlane M.;
*The mammalian CD4 homologue, APAF1, exists as two distinct forms in

RT human cells.";
 RL submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [17]
 RP SEQUENCE OF 1-138 FROM N.A. (ISOFORM 1/4/5).
 RA Won M., Lee J.-W., Ohn H.-H., Kim D.-U., Chung K.-S., Lee M.,
 RA Yoo H.-S.,
 RA "Cloning of variant Apatl.";
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RL [18]

RP INDUCTION BY E2F AND P53.
RX MEDLINE-21283226; Pubmed-11389439;
RA MOTORI M.C., HICKMAN E.S., DENCHI E.L., CAPRARA G., COLLI E.,
RA CECCONI F., MUELLER H., HELIN K.;
RT "Apaf-1 is a transcriptional target for E2F and p53."
RL Nat. Cell Biol. 3:552-558(2001).
RN 1701

RP X-RAY CRYSTALLOGRAPHY (1.3 ANGSTROMS) OF 1-97.
RX MEDLINE=20013059; PubMed=10543941;
RA Vaughn D.E., Rodriguez J., Lazebnik Y., Joshua-Tor L.;
RT "Crystal structure of Apaf-1 caspase recruitment domain: an alpha
RL helical Greek key fold for apoptotic signaling.";
RN J. Mol. Biol. 293:439-447.(1999).

RP STRUCTURE BY NMR OF 1-97.
RX MEDLINE-20047184; Pubmed-10578182;
RA Day C.L., Dupont C., Lackmann M., Vaux D.L., Hinds M.G.;
RT "Solution structure and mutagenesis of the caspase recruitment domain
RT (CARD) from Apaf-1";
RL Cell Death Differ. 6:1125-1132(1999).
RC -1- FUNCTION: Olfometric; Anst-1 mediates the nitrochryso-
-dendronant

CC -1- FUNCTION: Oligomeric Apaf-1 mediates the cytochrome c-dependent
CC autocatalytic activation of pro-caspase 9 (Apaf-3), leading to the
CC activation of caspase-3 and apoptosis. This activation requires
CC ATP.

CC -1- SUBUNIT: Monomer. Oligomerizes upon binding of cytochrome c and dAPP. Oligomeric Apat-1 and pro-caspase-9 bind to each other via their respective NH₂-terminal CARD domains and consecutively mature caspase-9 is released from the complex.

CC -1- SUBCELLULAR LOCATION: Cytoplasm;c.
CC -1- ALTERNATIVE PRODUCTS: 5 isoforms; 1/Apaf-1XL (shown here), 2/Apaf-
CC 1L, 3/Apaf-1S, 4/Apaf-1M and 5/Apaf-1XS; are produced by
CC alternative splicing

-I- TISSUE SPECIFICITY: Ubiquitous. Highest levels of expression in adult spleen and peripheral blood leukocytes, and in fetal brain, kidney and lung. Isoform 1 is expressed in heart, kidney and liver.

CC -1- INDUCTION: By E2F and p53 in apoptotic neurons.
CC -1- SIMILARITY: CONTAINS 1 CARD DOMAIN.
CC

CC -1- SIMILARITY: CONTAINS 13 WD REPEATS (TRP-ASP DOMAINS)

CC -1- SIMILARITY: CONTAINS 13 WD REPEATS (TRP-ASP DOMAINS)
CC -1- SIMILARITY: CONTAINS 1 NR-ABC DOMAIN

-1- CAUTION: Ref. 7 sequence differs from that shown due to a frameshift in position 109.

CC This swiss-prot entry is copyright. It is reproduced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR	EMBL; AF013263;	ACC51678.1;	-
DR	EMBL; AJ242400;	CAB55579.1;	-
DR	EMBL; AJ243004;	CAB5580.1;	-
DR	EMBL; AJ243005;	CAB55581.1;	-
DR	EMBL; AJ243006;	CAB55582.1;	-
DR	EMBL; AJ243007;	CAB55583.1;	-
DR	EMBL; AJ243008;	CAB55584.1;	-

EMBL: AJ243009; CAB55585.1; -
 EMBL: AJ243010; CAB55586.1; -
 EMBL: AJ243011; CAB55587.1; -
 EMBL: AJ243048; CAB55588.1; -
 EMBL: AJ243107; CAB55589.1; -
 EMBL: AF134397; CAB55590.1; -
 EMBL: AF149794; CAB55591.1; -
 EMBL: AB007873; CAB55592.1; -
 EMBL: AJ133643; CAB55593.1; -
 EMBL: AJ133644; CAB55594.1; -
 EMBL: AJ133645; CAB55595.1; -
 EMBL: AF248734; CAB55596.1; -
 PDB: 1CWS; 01-DEC-99.
 PDB: ICWS; 21-JAN-00.
 PDB: 602233; -
 InterPro: IPR001315; CARD.
 InterPro: IPR002182; NB-ARC.
 InterPro: IPR001680; WD40.
 Pfam: PF00619; CARD; 1.
 Pfam: PF00931; NB-ARC; 1.
 Pfam: PF00400; WD40; 1.
 PRINTS: PR00320; GPROTEINRPT.
 SMART: SM00320; WD40; 10.
 PROSITE: PS50209; CARD; 1.
 PROSITE: PS00678; WD_REPEATS_1; 4.
 PROSITE: PS50082; WD_REPEATS_2; 9.
 PROSITE: PS50294; WD_REPEATS_REGION; 1.
 Apoptosis: ATP-binding; Repeat; WD repeat; Alternative splicing;
 3D-structure.
 DOMAIN 1 90
 NP_BIND 104 415
 REPEAT 154 161
 REPEAT 613 652
 REPEAT 655 694
 REPEAT 697 738
 REPEAT 741 780
 REPEAT 796 836
 REPEAT 838 877
 REPEAT 880 919
 REPEAT 959 998
 REPEAT 1001 1040
 REPEAT 1042 1080
 REPEAT 1083 1122
 REPEAT 1125 1164
 REPEAT 1175 1212
 DOMAIN 95 98
 VARSPLIC 99 109
 VARSPLIC 575 575
 VARSPLIC 824 866
 VARSPLIC 1113 1154
 CONFLICT 108 108
 CONFLICT 134 134
 CONFLICT 145 145
 CONFLICT 161 161
 CONFLICT 370 370
 CONFLICT 383 383
 CONFLICT 544 544
 CONFLICT 580 580
 CONFLICT 608 608
 CONFLICT 620 620
 CONFLICT 639 639
 CONFLICT 708 708
 CONFLICT 742 742
 CONFLICT 746 746
 CONFLICT 757 757
 CONFLICT 795 795
 CONFLICT 798 798
 CONFLICT 825 825
 CONFLICT 871 871
 CONFLICT 876 876
 CONFLICT 949 949
 CARD.
 NB-ARC.
 ATP (POTENTIAL).
 WD 1.
 WD 2.
 WD 3.
 WD 4.
 WD 5.
 WD 6.
 WD 7.
 WD 8.
 WD 9.
 WD 10.
 WD 11.
 WD 12.
 WD 13.
 POLY-SER.
 MISSING (IN ISOFORM 2 AND ISOFORM 3).
 E -> ETLGESSK (IN ISOFORM 5).
 MISSING (IN ISOFORM 3, ISOFORM 4 AND
 ISOFORM 5).
 MISSING (IN ISOFORM 5).
 S -> P (IN REF. 7).
 N -> S (IN REF. 7).
 G -> C (IN REF. 2; CAB55587).
 S -> F (IN REF. 2; CAB55586).
 I -> T (IN REF. 2; CAB55581).
 Y -> H (IN REF. 2; CAB55586).
 F -> T (IN REF. 2; CAB55584).
 A -> L (IN REF. 2; CAB55580).
 R -> C (IN REF. 2; CAB55585).
 H -> R (IN REF. 2; CAB55579).
 L -> F (IN REF. 2; CAB55583).
 T -> A (IN REF. 2; CAB55579).
 H -> R (IN REF. 2; CAB55583).
 V -> P (IN REF. 2; CAB55584).
 L -> P (IN REF. 2; CAB55582).
 E -> G (IN REF. 2; CAB55581).
 D -> A (IN REF. 2; CAB55583).
 S -> L (IN REF. 2; CAB55587).
 A -> T (IN REF. 2; CAB55581).
 I -> V (IN REF. 2; CAB55585).

Query Match 2.9%; Score 104.5; DB 1; Length 1248;
 Best Local Similarity 19.1%; Pred. No. 2.5; Mismatches 188; Indels 221; Gaps 27;
 Matches 112; Conservative 64; Mismatches 188; Indels 221; Gaps 27;
 QY 79 DCALRRVYVHLAVGH-----WAEPEAGKSDSELOE--SRNASLQA-QVYLSQAV-- 128
 DB 537 DCAVSENFQEFSLNGHLGRPFNIVQGLCEPEISYVQAQKQAKQEDNMLYLE 596
 QY 129 -----PIKCALLEVQVADLV-----OPQSVGSAVFD---CFEASLIG----- 164
 DB 597 WINKKNITNLSRLVVRPHPDVAVHACFSESDGRIASGADKTLQVFAETGEKLEIKAH 656
 QY 165 -----AEVQIWS-----YTKPKYQKELNITQQLPDGDNVIL 195
 DB 657 EDEVLCCAFSTDDRIYACSVNCKVAINMSMTGELVHTYDHSQVNCHEFTNSHHLL 716
 QY 196 TLVDSEEDFSFLYLRVPYDALKSLMYKNLTGPONITLNTDLPVLCIQVWSLEPDS 255
 DB 717 ATGSS-----DCFLKMDLNQKCRNTMFGHTN----- 744
 QY 256 RVECFCEPREDGCAHRLNHLIARLVLSPGVWOLDAPCLPGKVTLCQAPDQSPQPLVP 315
 DB 745 SVNHCRESPPD-----KILASGADGTLKL-WDATSANEKSI-- 781
 QY 316 PVPKNAIVNPPRODFQVAGHPNLCVQVSTWEVQLQACSW-ADS-----LGP 362
 DB 782 NVKQFLNLEDPQEDM-----EVYKCCSWSDGARIMAAKNKIFLED 825
 QY 363 FKDMLEVMKGTGLNNT-SVCALEPSCGTPPLPSMASTRARLRELLQDFRSHOCMLAN 421
 DB 826 IHTSGLEIHTGHSTLOYODFSPQNHIAVAL-----SQCVELMN 868
 QY 422 DDWNGSLMACPMKXTHRRWVWLACLLIAALFFELLKKDKRRKARSGRTALLHSA 481
 DB 869 TDSRSKVDKRG-----HLSVW-----HG-----VMSFP 892
 QY 482 DGAGYERLVGALSALSGMPLRVAVDMSRRELSANGALAMFHQRRRIIOGCVVILF 541
 DB 893 DGSSF-----LTSSDDO-----TIRLMETKRYCKNSAV-MLKQEDVVFQENEMVVL- 938
 QY 542 SPAVAQCQOQMLQOTVPEPQHDAL--AAMLSVLPDPLQGRANG 584
 DB 939 ---AVDHIRR-LQILINGTGQIDVLTQAVSCCISPHQYIARF 979
 RESULT 5
 PGBM_MOUSE STANDARD; PRT; 3707 AA.
 ID PGBM_MOUSE
 AC 005793;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, last sequence update)
 DT 01-MAR-2002 (Rel. 41, last annotation update)
 DE Basement membrane-specific heparan sulfate proteoglycan core
 protein precursor (HSPC) (Pellecan) (PDCP).
 GN HSPC2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Melanoma;
 RA MEDLINE=92078153; PubMed=1744087;
 RA Noonan D.M., Fulle A., Valente P., Cai S., Horigan E., Sasaki M.,
 Yamada Y., Hassell J.R.,
 RT *The complete sequence of perlecan, a basement membrane heparan
 RT sulfate proteoglycan, reveals extensive similarity with laminin A
 RT chain, low density lipoprotein-receptor, and the neural cell adhesion
 RT molecule*;
 RL J. Biol. Chem. 266:22939-22947 (1991).
 RN [2]
 RP SEQUENCE OF 940-1601 AND 1870-2600 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=89034110; PubMed=2972708;

Noonan D.M., Horigan E.A., Ledbetter S.R., Vogeli G., Sasaki M., Yamada Y., Hassell J.R.;

Identification of cDNA clones encoding different domains of the basement membrane heparan sulfate proteoglycan.*;

J. Biol. Chem. 263:16379-16387(1988).

-1- FUNCTION: THIS PROTEIN IS AN INTEGRAL COMPONENT OF BASEMENT MEMBRANES. IT IS RESPONSIBLE FOR THE FIXED NEGATIVE ELECTROSTATIC CHARGE AND IS INVOLVED IN THE CHARGE-SELECTIVE ULTRAFILTRATION PROPERTIES. IT INTERACTS WITH OTHER BASEMENT MEMBRANE COMPONENTS SUCH AS LAMININ AND COLLAGEN TYPE IV AND SERVES AS AN ATTACHMENT SUBSTRATE FOR CELLS.

-1- SUBUNIT: PURIFIED PERLECAN HAS A STRONG TENDENCY TO AGGREGATE IN DIMERS OR STELLATE STRUCTURES.

-1- SUBCELLULAR LOCATION: Extracellular.

-1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES.

-1- PTM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED AND O-LINKED OLIGOSACCHARIDES.

-1- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.

-1- SIMILARITY: CONTAINS 10.5 LAMININ EGF-LIKE DOMAINS.

-1- SIMILARITY: CONTAINS 15 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

-1- SIMILARITY: CONTAINS 1 SEA DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL: M77174; AAA39911.1; -

EMBL: J04054; AAA39899.1; -

EMBL: J04055; AAA39912.1; -

HSSP: Q07954; 1CR8.

MGD: MG1:96257; HSP92.

InterPro: IPR000561; EGF-like.

InterPro: IPR000742; EGF_2.

InterPro: IPR001438; EGF_11.

InterPro: IPR003006; Ig_MHC.

InterPro: IPR003598; Ig_C2.

InterPro: IPR002172; LDL_recept_A.

InterPro: IPR000034; Laminin_B.

InterPro: IPR002049; Laminin_EGF.

InterPro: IPR001791; Laminin_G.

InterPro: IPR000082; SEA.

Pfam: PF00047; Ig_15.

Pfam: PF00052; Laminin_B; 3.

Pfam: PF00053; Laminin_EGF; 8.

Pfam: PF00054; Laminin_G; 3.

Pfam: PF00057; LDL_recept_A; 4.

Pfam: PF01390; SEA; 1.

PRINTS: PR00010; EGFBL00D.

ProDom: PD003031; Laminin_B; 3.

SMART: SM00180; EGF_Lam; 7.

SMART: SM00001; EGF_Like; 6.

SMART: SM00408; IGC2; 14.

SMART: SM00192; LDLA; 4.

SMART: SM00281; Lamb; 3.

SMART: SM00282; Lamb; 3.

SMART: SM00200; SEA; 1.

PROSITE: PS00022; EGF_1; 8.

PROSITE: PS01186; EGF_2; 5.

PROSITE: PS01248; LAMININ_TYPE_EGF; 11.

PROSITE: PS50025; LAM_G_DOMAIN; 3.

PROSITE: PS01209; LDLRA_1; 4.

PROSITE: PS50068; LDLRA_2; 4.

PROSITE: PS50024; SEA; 1.

Signal: Basement membrane; Proteoglycan; Repeat; Glycoprotein; Heparan sulfate; Laminin EGF-like domain; Immunoglobulin domain; Extracellular matrix; EGF-like domain.

FT	SIGNAL	1	21	POTENTIAL.
FT	CHAIN	22	3707	BASEMENT MEMBRANE-SPECIFIC HEPARAN SULFATE PROTEOGLYCAN CORE PROTEIN.
FT	DOMAIN	80	194	SEA.
FT	DOMAIN	195	234	LDL-RECEPTOR CLASS A 1.
FT	DOMAIN	281	319	LDL-RECEPTOR CLASS A 2.
FT	DOMAIN	320	359	LDL-RECEPTOR CLASS A 3.
FT	DOMAIN	360	403	LDL-RECEPTOR CLASS A 4.
FT	DOMAIN	404	504	IG-LIKE C2-TYPE DOMAIN 1.
FT	DOMAIN	521	530	LAMININ EGF-LIKE 1 (N-TERMINAL).
FT	DOMAIN	531	730	LAMININ EGF-LIKE 1 (N-TERMINAL).
FT	DOMAIN	731	763	LAMININ EGF-LIKE 1 (C-TERMINAL).
FT	DOMAIN	764	813	LAMININ EGF-LIKE 2.
FT	DOMAIN	814	871	LAMININ EGF-LIKE 3.
FT	DOMAIN	879	923	LAMININ EGF-LIKE 4 (INCOMPLETE).
FT	DOMAIN	924	933	LAMININ EGF-LIKE 5 (N-TERMINAL).
FT	DOMAIN	934	1125	LAMININ EGF-LIKE 5 (C-TERMINAL).
FT	DOMAIN	1126	1158	LAMININ EGF-LIKE 5 (C-TERMINAL).
FT	DOMAIN	1159	1208	LAMININ EGF-LIKE 6.
FT	DOMAIN	1209	1265	LAMININ EGF-LIKE 7.
FT	DOMAIN	1275	1324	LAMININ EGF-LIKE 8.
FT	DOMAIN	1325	1334	LAMININ EGF-LIKE 9 (N-TERMINAL).
FT	DOMAIN	1335	1529	LAMININ EGF-LIKE 9 (N-TERMINAL).
FT	DOMAIN	1530	1562	LAMININ EGF-LIKE 9 (C-TERMINAL).
FT	DOMAIN	1563	1612	LAMININ EGF-LIKE 10.
FT	DOMAIN	1613	1670	LAMININ EGF-LIKE 11.
FT	DOMAIN	1671	1771	IG-LIKE C2-TYPE DOMAIN 2.
FT	DOMAIN	1772	1865	IG-LIKE C2-TYPE DOMAIN 3.
FT	DOMAIN	1866	1954	IG-LIKE C2-TYPE DOMAIN 4.
FT	DOMAIN	1955	2049	IG-LIKE C2-TYPE DOMAIN 5.
FT	DOMAIN	2050	2148	IG-LIKE C2-TYPE DOMAIN 6.
FT	DOMAIN	2149	2244	IG-LIKE C2-TYPE DOMAIN 7.
FT	DOMAIN	2245	2343	IG-LIKE C2-TYPE DOMAIN 8.
FT	DOMAIN	2344	2436	IG-LIKE C2-TYPE DOMAIN 9.
FT	DOMAIN	2437	2532	IG-LIKE C2-TYPE DOMAIN 10.
FT	DOMAIN	2533	2619	IG-LIKE C2-TYPE DOMAIN 11.
FT	DOMAIN	2620	2720	IG-LIKE C2-TYPE DOMAIN 12.
FT	DOMAIN	2721	2809	IG-LIKE C2-TYPE DOMAIN 13.
FT	DOMAIN	2810	2895	IG-LIKE C2-TYPE DOMAIN 14.
FT	DOMAIN	2896	2980	IG-LIKE C2-TYPE DOMAIN 15.
FT	DOMAIN	2984	3162	LAMININ G-LIKE 1.
FT	DOMAIN	3163	3241	EGF-LIKE.
FT	DOMAIN	3245	3425	LAMININ G-LIKE 2.
FT	DOMAIN	3518	3705	LAMININ G-LIKE 3.
FT	SITE	65	67	HEPARAN SULFATE (POTENTIAL).
FT	SITE	71	73	HEPARAN SULFATE (POTENTIAL).
FT	SITE	76	78	HEPARAN SULFATE (POTENTIAL).
FT	SITE	3615	3617	MEDIATES MOTOR NEURON ATTACHMENT (POTENTIAL).
FT	DISULFID	199	212	BY SIMILARITY.
FT	DISULFID	206	225	BY SIMILARITY.
FT	DISULFID	219	234	BY SIMILARITY.
FT	DISULFID	285	297	BY SIMILARITY.
FT	DISULFID	292	310	BY SIMILARITY.
FT	DISULFID	304	319	BY SIMILARITY.
FT	DISULFID	325	337	BY SIMILARITY.
FT	DISULFID	332	350	BY SIMILARITY.
FT	DISULFID	344	359	BY SIMILARITY.
FT	DISULFID	368	381	BY SIMILARITY.
FT	DISULFID	375	394	BY SIMILARITY.
FT	DISULFID	388	403	BY SIMILARITY.
FT	DISULFID	428	479	BY SIMILARITY.
FT	DISULFID	764	773	BY SIMILARITY.
FT	DISULFID	766	780	BY SIMILARITY.
FT	DISULFID	783	792	BY SIMILARITY.
FT	DISULFID	795	811	BY SIMILARITY.
FT	DISULFID	814	829	BY SIMILARITY.
FT	DISULFID	816	839	BY SIMILARITY.
FT	DISULFID	854	851	BY SIMILARITY.
FT	DISULFID	854	869	BY SIMILARITY.
FT	DISULFID	1159	1168	BY SIMILARITY.
FT	DISULFID	1161	1175	BY SIMILARITY.
FT	DISULFID	1178	1187	BY SIMILARITY.

DISULFID	1190	1206	BY SIMILARITY.
DISULFID	1209	1224	BY SIMILARITY.
DISULFID	1211	1234	BY SIMILARITY.
DISULFID	1237	1246	BY SIMILARITY.
DISULFID	1249	1263	BY SIMILARITY.
DISULFID	1275	1287	BY SIMILARITY.
DISULFID	1277	1293	BY SIMILARITY.
DISULFID	1295	1304	BY SIMILARITY.
DISULFID	1307	1322	BY SIMILARITY.
DISULFID	1563	1572	BY SIMILARITY.
DISULFID	1565	1579	BY SIMILARITY.
DISULFID	1582	1591	BY SIMILARITY.
DISULFID	1594	1610	BY SIMILARITY.
DISULFID	1613	1628	BY SIMILARITY.
DISULFID	1641	1650	BY SIMILARITY.
DISULFID	1653	1668	BY SIMILARITY.
DISULFID	1792	1839	BY SIMILARITY.
DISULFID	1886	1932	BY SIMILARITY.
DISULFID	1976	2021	BY SIMILARITY.
DISULFID	2073	2118	BY SIMILARITY.
DISULFID	2170	2215	BY SIMILARITY.
DISULFID	2268	2313	BY SIMILARITY.
DISULFID	2365	2413	BY SIMILARITY.
DISULFID	2456	2506	BY SIMILARITY.
DISULFID	2554	2599	BY SIMILARITY.
DISULFID	2641	2686	BY SIMILARITY.
DISULFID	2831	2876	BY SIMILARITY.

Query Match
Best Local Similarity 2.98; Score 103.5; DB 1; Length 3707;
Matches 87; Conservative 44; Mismatches 167; Indels 127; Gaps 19;

26	EPQDPTARCSLG-LSCHW--DGDVLCPLGSLQ-----APGVLVPTROTETELVLR----	73
326	ERNRFA-CEHGKALKLMKRCDDPDCEDRTDANC SVKPGEGCGPHFCVSTNRCTIPA	384
74	---SPKTDALRVVVVHLVHGMARPEAKSDSELOESNNAQLQAQVLSFGQVYPI	130
385	SFHCDESDCPDR-----SDEGCMRPQVYTPPQOSIQ-----	418
131	ARCALLEVQVPAIDLVPGQSVGSAVFDFEASLSGAELVQIWSYTKPRYKELNTQDLPDG	190
419	-----SRGQVY--TFTCVATGVPTPIIMWRLMWHGHPAPRVTMTSEGG	460
191	DNVLTLDVSEEDDFSL-----YLRVPDAKSLMYKNIKTGQN--TILNHTDLPV	241
461	RGTLLTRDYKADOGAATTCCEAMNSRGAVFGIPDGVLELVPQRGPPCDGHFYLEDSASCLP	520
242	CLCIQWMSLEPDERVEFCFEDEPGAHRLNHLIARLVLSFGVMDLAPCCLPG-KVTL	300
521	CECFGVTVNCQSSLR-----FRD-----QIRL-----SFDQPNDRKGVAVTM	557
301	CMDAPQSPQPLVPPVPOKNAVYNEP-QDFOLVAGHPNLVQVSTWKEVQLOACSMADS	359
558	-----PSQPPVPLPSTQLQIDPLAQEFQLVDSRFLVHADFMALPKQFLGNKYDS	609
360	LGPFKMDMLVEKKTGNTNNSVCLPE-----SCTPL-PSMASTRAA	401
610	YCGFLAKYKRYELARGM-----LEPVQKPDVILVAGGYRLHSRGHTPHPGTLNROV	662
402	RLGEE 406	
663	OLSEE 667	

RESULT 6
NAL2_HUMAN STANDARD; PRT: 1062 AA.
O9NWX2: O9BNV5: O9HG66: O9HAV9: O9NMK3;
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)

DE NACHT-, LRR- and PYD-containing protein 2 (Nucleotide-binding site
protein 1).
GN NALP2 OR NBS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE-21169419; PubMed-11270363;
RA Bertin J., Distefano P.S.;
RT "The PYRIN domain: a novel motif found in apoptosis and inflammation
proteins.";
RL Cell Death Differ. 7:1273-1274(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE-21148093; PubMed-11250163;
RA Martinon F., Hoffmann K., Tschopp J.;
RT "The pyrin domain: a possible member of the death domain-fold family
implicated in apoptosis and inflammation.";
RL Curr. Biol. 11:R118-R120(2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE-Kidney epithelium, and Colon;
RA Matanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Ohayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isegal T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE-Lung, and Placenta;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: May be implicated in apoptosis (By similarity).
CC -1- COFACTOR: Binds ATP (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC -1- SIMILARITY: CONTAINS 1 DAPIN DOMAIN.
CC -1- SIMILARITY: CONTAINS 9 NACHT DOMAIN.
CC -1- SIMILARITY: CONTAINS 9 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AF298547; AAC15253.1; ALT_INIT.
CC EMBL: AF310106; AAG30289.1; -
CC EMBL: AK000517; BAA91223.1; -
CC EMBL: AK025952; BAB15293.1; -
CC EMBL: AK000784; BAB1377.1; ALT_INIT.
CC EMBL: BC003592; AAH03592.1; -
CC EMBL: BC001039; AAH01039.1; -
CC InterPro: IPR000767; Disease_resist.
CC InterPro: IPR003592; LRR_out.
CC InterPro: IPR004020; PAAD_DAPIN.
CC PRINTS: PR00364; DISEASEREST.
CC SMART: SM00370; LRR: 7.
CC PROSITE: PS50824; DAPIN: 1.
CC PROSITE: PS50837; NACHT: 1.
CC Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
KW Alternative splicing.
FT DOMAIN 1 94 DAPIN.
FT DOMAIN 207 526 NACHT.
FT REPEAT 467 491 LRR 1.
FT REPEAT 622 645 LRR 2.
FT REPEAT 754 777 LRR 3.
FT REPEAT 810 832 LRR 4.

REPEAT	839	862	LRR 5.
REPEAT	867	890	LRR 6.
REPEAT	924	947	LRR 7.
REPEAT	981	1005	LRR 8.
REPEAT	1010	1033	LRR 9.
NP_BIND	213	220	ATP (POTENTIAL).
DOMAIN	518	523	POLY-GLU.
VARSPLIC	133	154	MISSING (IN ISOFORM 2).
CONFLICT	1	1	M -> V (IN REF. 1).
CONFLICT	35	35	L -> P (IN REF. 1).
CONFLICT	304	304	I -> S (IN REF. 3; BAB15293).
CONFLICT	364	364	R -> K (IN REF. 1).
CONFLICT	980	980	MISSING (IN REF. 1).
CONFLICT	1052	1052	A -> E (IN REF. 4).
SEQUENCE	1062 AA;	120514 MM;	4DB0F6E9C2BCBA7 CRC64;

Query Match 2.88; Score 100.5; DB 1; Length 1062;
Dist Local Similarity 22.38; Pred. No. 4.2;
Matches 126; Conservative 62; Mismatches 177; Indels 199; Gaps 34;

46	VLCPLPSLASAPGVLPVRL-OEELYLAR-C---POKIDCALRYVYVHVAHGMAEP	99
403	ICTTTLKLOMKEGDEPVTCLTGLTRFLCISRPQGAOLGALRTLSLLAQAQGLMAOT	462
100	-----EAGKSDSEL-----OESRNASIAQVYVLFQAYPIARCALLEVOYE	141
463	SVLHREDLERLGVOESDLRLFLDGDILKQDNRVSKGCSFPHLSFOOFLTALPTLKEEE	522
142	ADLVOPGQSVGSAVFDFEASLAE-----VOIWSYT---KPRYOKELNLT---QOL	187
523	EDRDGHTWIG-----DVOKILSGVERLRNPDLIQAGYSGFLGNAEKRAELETFGCBS	578
188	PGSDNVLLTLDS-----EODSFLLYLARVPALASLM--YKMLTGPONTLNH	236
579	PIKQELRLKCDISCKGSHSTVTDLQELGLCELYSEELVEVMAOFKEI---SLHLNA	634
237	TDLPV-----CICIQVWSL-----EPDSERVE-----FCPPREDP--	266
635	VDVVSFVCVKHNRNLOKMSLOYIKENLPENVTASBSAE-VERQSDODHMLPFTDILCS	693
267	--GAINRLMHAIA-RLNVLSPGWOQLAPCCLPGRVTLQW--APDQSPCOPLVPPQKNA	322
694	IFGSKNDLWGIAINDSEFSLASIVRI-----LCEQIASDCHIQORV--FKNI	728
323	TVNEQDQFLVAGHPMLCY-----QVSTWENVO-----L	351
739	S---PAD-----AHRNLCIALRGKHKTVTYTLTGQNDODDMEPALCEVLRNHPRCNLT	790
352	QACS-----WAD-SLG-----PFKDDMLVE-----MKTGLNNT	379
791	VSCSAFTQOQADSLALEVNSILTCVNLSDNEILLDEBKKLYTTLRHKRCFLQRLSLEN-	849
380	SVCALEPSCGCTPLPS-----MASTRARLGEELL-ODFRSHOC-----MOLAND	422
850	--CHLTENACKDLAALVAVSRRETLHLCLAKNPIGNTGKFLCEGLRARECKQTITLVNMC	907
423	DNMGSLMCPMKYIHRNVLVWL	446
908	D-ITSDGCCDITKLLQEKSLTCL	930

RESULT 7
DCC_HUMAN STANDARD; PRT; 1447 AA.
P43146;
01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
Tumor suppressor protein DCC precursor (Colorectal cancer suppressor).
DCC.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE-95011532; PubMed-7926722;
RA Hedrick L., Cho K.R., Fearon E.R., Wu T.-C., Kinzler K.W.,
RA Vogelstein B.;
RT "The DCC gene product in cellular differentiation and colorectal
RL tumorigenesis.";
RN Genes Dev. 8:1174-1183(1994).
RN [2]
RN SEQUENCE OF 1-750 FROM N.A.
RX MEDLINE-90100559; PubMed-2294591;
RA Fearon E.R., Cho K.R., Nigro J.M., Kern S.E., Simons J.W.,
RA Ruppert J.M., Hamilton S.R., Preisinger A.C., Thomas G., Kinzler K.W.,
RA Vogelstein B.;
RT "Identification of a chromosome 18q gene that is altered in
RL colorectal cancers.";
RN Science 247:49-56(1990).
RN [3]
RN SEQUENCE OF 107-472 FROM N.A. (SCRAMBLED EXONS).
RX MEDLINE-91121517; PubMed-1991322;
RA Nigro J.M., Cho K.R., Fearon E.R., Kern S.E., Ruppert J.M.,
RA Oliner J.D., Kinzler K.W., Vogelstein B.;
RT "Scrambled exons.";
RL Cell 64:607-613(1991).
RN [4]
RN GENE STRUCTURE, AND VARIANTS CARCINOMA HIS-1375.
RX MEDLINE-94245241; PubMed-8188295;
RA Cho K.R., Oliner J.D., Simons J.W., Hedrick L., Fearon E.R.,
RA Preisinger A.C., Hedge P., Silverman G.A., Vogelstein B.;
RT "The DCC gene: structural analysis and mutations in colorectal
RL carcinomas.";
RN Genomics 19:525-531(1994).
RN [5]
RN VARIANT CARCINOMA THR-168, AND VARIANT GLY-201.
RX MEDLINE-94243823; PubMed-8187090;
RA Miyake S., Nagai K., Yoshino K., Oto M., Endo M., Yuasa Y.;
RT "Point mutations and allelic deletion of tumor suppressor gene DCC in
RT human esophageal squamous cell carcinomas and their relation to
RL metastasis.";
RL Cancer Res. 54:3007-3010(1994).
CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: FOUND IN AXONS OF THE CENTRAL AND PERIPHERAL
CC NERVOUS SYSTEM AND IN DIFFERENTIATED CELL TYPES OF THE INTESTINE.
CC -1- DISEASE: COLORECTAL TUMORS THAT LOST THEIR CAPACITY TO
CC DIFFERENTIATE INTO MUCUS PRODUCING CELLS UNIFORMITY LACK DCC
CC EXPRESSION. INACTIVATION OF DCC DUE TO ALLELIC DELETION AND/OR
CC POINT MUTATIONS MAY CAUSE BOTH LYMPHATIC AND HEMATOGENOUS
CC METASTASIS OF OESOPHAGEAL SQUAMOUS CELL CARCINOMAS.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X76132; CAA53735.1; -
CC EMBL; M32292; AAA35751.1; -
CC EMBL; M32286; AAA52174.1; -
CC EMBL; M32288; AAA52175.1; ALT_SEQ.
CC EMBL; M32290; AAA52176.1; -
CC EMBL; M32296; AAA52177.1; -
CC EMBL; M63700; AAA52178.1; -
CC EMBL; M63702; AAA52179.1; -
CC EMBL; M63718; AAA52180.1; -

Db	188	RYVYLFESGALQISLQPGDIGITCYCSAPNPNASSRTGNEAEVNLISDPGLHQLYFQRP	247
Oy	276	-----	279
Db	248	NYVAIEGKDAVLEECVSGYPPEPFTWLGEVYIQLSKKYSLLGSGNLLISNTVDDSGM	307
Oy	276	-----	318
Db	308	YTCVVTYKKNENISASAELTYLVP-PWFLNHPNSL-----YAYESMDIEFFECTVSGKPVPT	361
Oy	319	---OKNATVNEPODEFQVAVGHPNLCVQVSTWEKVLQACASMDNISGPEFDD-----MLTVE	371
Db	362	VMMKMGDVIPSPDYFQIVGSSNLRI-----LGVKSDSGEFTQCVAR	403
Oy	372	MKTGLNNTSVCALEPSCGP-----LPSMASTRARALGEELLQDFRSHOCMOL	419
Db	404	NEAGNQSOTSQOLIVPKPAIFSSSVLPAPRDVVPVLYVSSEFVRLS-----	448
Oy	420	WMDNNGSILMACMDKIYHRWVLYWLACILLAALFFFLILKDKRRKARS-RFALL	478
Db	449	-----WRPPE-----AKGNIQTFYV	465
Oy	479	HSADGAGYERLYGALASALSQMPR-----VAVDLSRRELSAHGALAFHHQ	526
Db	466	FREGDNREBALMTTQPGSLOLYTGNLKPREAMYTFRRVAYANMGGESS-----	514
Oy	527	RRRIQEGGVILLFSPAVALCOQOMIQ-----LQTVRPGPHDALAMLSCVLPFIQ	579
Db	515	-----OPIKVA-TQPEIQVPGPVENIQAIVSTFSPTSLITWEP---PAYAN	555
Oy	580	GR-----ATGRYGVYFPG	594
Db	556	GPVQGRFLPCTEYVSTGEKONIEVDGL	581
RESULT 8			
ABP2_HUMAN			
ID	ABP2_HUMAN	STANDARD:	PRT: 2647 AA.
AC	P21333:		
DT	01-MAY-1991	(Rel. 18, Created)	
DT	01-MAY-1991	(Rel. 18, last sequence update)	
DT	16-OCT-2001	(Rel. 40, last annotation update)	
DE	Endothelial actin-binding protein (ABP-280) (Nonmuscle filamin)		
DE	(Filamin 1) (Alpha-filamin).		
GN	FLNI OR FLN.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RX	MEDLINE=90361737;	PubMed=2391361;	
RA	Goelin J.B., Yamai R., Egan S., Stewart W., Stossel T.P.,		
RA	Kwiatkowski D.J., Hartwig J.H.;		
RT	Human endothelial actin-binding protein (ABP-280, nonmuscle		
RT	filamin): a molecular leaf spring.		
RL	J. Cell Biol. 111:1089-1105(1990).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=96311563;	PubMed=8733135;	
RA	Zhu E.Y., Zollo M., Mazzarella R.A., Ciccodicola A., Chen C.-N.,		
RA	Chou L., Heider C., Burrough F.W., Ripetto M., Schlessinger D.,		
RA	D'Urso M.;		
RT	Long-range sequence analysis in Xq28: thirteen known and six		
RT	candidate genes in 219.4 kb of high GC DNA between the RCP/GCP and		
RT	Hum. Mol. Genet. 5:659-668(1996).		
RL	[3]		
RN	[3]		
RP	SEQUENCE OF 1658-1772 FROM N.A.		
RX	MEDLINE=93357748;	PubMed=7689010;	
RA	Mestrini E., Patrosso C., Mancini M., Rivella S., Rocchi M.,		
RA	Repetto M., Villa A., Fratini A., Zoppe M., Vezoni P.		

Tomolo D.: Mapping of two genes encoding isoforms of the actin binding protein Abp-280, a dystrophin like protein, to Xq28 and to chromosome 7. Hum. Mol. Genet. 2:761-766(1993).

- FUNCTION: PROMOTES ORTHOGONAL BRANCHING OF ACTIN FILAMENTS AND LINKS ACTIN FILAMENTS TO MEMBRANE GLYCOPROTEINS.

- SUBUNIT: HOMODIMER. INTERACTS WITH CYPSP.

- SUBCELLULAR LOCATION: PERIPHERAL CYTOPLASM.

- TISSUE SPECIFICITY: UBIQUITOUS.

- PTM: PHOSPHORYLATION EXTENT CHANGES IN RESPONSE TO CELL ACTIVATION.

- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN, ABP-120, ABP-180, OR BETA-FODRIN).

- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.

- SIMILARITY: CONTAINS 24 FILAMIN REPEATS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X53416; CAA37495.1; -
EMBL; L44140; AA92644.1; -
EMBL; X70082; CAA49687.1; -
EMBL; X70085; CAA49690.1; -
PIR; A37098; A37098.
HSSP; P13466; 1KSR.
MIM; 300017; -
InterPro; IPR001589; Actinin_act_bind.
InterPro; IPR001715; Calponin_hom.
InterPro; IPR001298; Filamin.
Pfam; PF00307; CH; 2.
Pfam; PF00630; Filamin; 24.
SMART; SM00033; CH; 2.
PROSITE; PS00019; ACTININ_1; 1.
PROSITE; PS00020; ACTININ_2; 1.
PROSITE; PS00021; ACTININ_3; 1.
PROSITE; PS50194; FILAMIN_REPEAT; 24.
Actin-binding; Phosphorylation; Repeat.
DOMAIN 1 274
DOMAIN 2 149
DOMAIN 3 166
DOMAIN 4 276
DOMAIN 5 376
DOMAIN 6 474
DOMAIN 7 570
DOMAIN 8 663
DOMAIN 9 763
DOMAIN 10 866
DOMAIN 11 965
DOMAIN 12 1061
DOMAIN 13 1154
DOMAIN 14 1249
DOMAIN 15 1349
DOMAIN 16 1442
DOMAIN 17 1539
DOMAIN 18 1636
DOMAIN 19 1740
DOMAIN 20 1778
DOMAIN 21 1860
DOMAIN 22 1950
DOMAIN 23 2039
DOMAIN 24 2131
DOMAIN 25 2230
DOMAIN 26 2335
DOMAIN 27 2420
DOMAIN 28 2516
DOMAIN 29 2551
DOMAIN 30 2646
DOMAIN 31 2647

FT	SITE	1761	1762	CLEAVAGE (BY CALPAIN).
FT	MOD_RES	1	1	BLOCKED.
FT	CONFLICT	1772	1772	A -> G (IN REF. 3).
FT	CONFLICT	2634	2634	H -> D (IN REF. 2).
SO	SEQUENCE	2647 AA;	280759 MM;	6C1A07041DFA3D42 CRC64;

Query Match

Best Local Similarity 20.68; Score 99.5; DB 1; Length 2647; Matches 145; Conservative 76; Mismatches 197; Indels 285; Gaps 42;

OY

122

VLSFOAYPIARCLLLEVOVADLVOP-----GOSVGSANFDCFEASLGAE-----VOIMSY 172

DB

537

VGFEETPM-----VGYITVITWGGONIGRSP--FEVVGTECGNOKVRAMG- 583

OY

173

TKPRYQELMLTQQLPBDGNVLLTLDVSEODEFLLYLRVPDAKSLMYKNLTGPONI 232

DB

584

-----PGLGEGV-----VGKSADE-----VEALIGDVGITGLF-SVEGPSOA 619

OY

233

TLNHTDVLPLCL-IQWSLEP-----DSERVECPREDPGAHRIIMHARLVLS 282

DB

620

KICEDCKGSGSCDVRYPQAGEYAVHYLCNSEDIRLSFPMADIRAPDPFPRVARG 679

OY

283

PGWQGLAPCCLP-----GKVTLCWQAPDQSPQ-----PLVPPV 317

DB

680

PLEKTVGAVNKRPAEFTYDANKHGRAPLRVQVODNEGCEVEALVKDNGNGTSCSYPRK 739

OY

318

PORNAVY-----NEPODFOLVAG--HPNLGVGVSTWEKVLQACSNADSLGPFKDDM 367

DB

740

PKHTHNAVSGVSIINSPFRVNVAGASHPN-----KRYV-----YGP----- 777

OY

368

LIVEMKGLNNTSVCALEPS-----GCTPLPSMASTRARLGEELLQDF 411

DB

778

-----GVAKTGLKAHEPTFTVDCAEAGQDVSIGICAPGVGPAPADIDFDIIR-- 828

OY

412

RSHQCMQMLNDNMGSLMACPMQYIHRRWYLVWLACLLAALEFFLLKDR----- 465

DB

829

-----NDNPTFY-----KTPR-----GAGSYTIMVLFADQAPPTSP 861

OY

466

-----RKAARG--SRTALLL--H--SADGAGYERL----- 489

DB

862

IRKVEPSHDSKVRKADGPOLSRIGVELEKPRHFTVNNAAAGKGLDVOFSGLTKGDVVR 921

OY

490

-----VGAIAS--ALSOMPLRAVAV-----DIMSRELSAHG 518

DB

922

DVDIIDHNDNTYVKKYTPVOQGPGVGNVNTYGGDPIPKSPFSVAVSPSLD-SKIVSGIG 980

OY

519

ALA-----WEHQRRIILQEGGVILLFSRA-AVAGCCQMLQLOTVGEGPHDALAAMLS 571

DB

981

EKYDVGKDOEFYVSKAGAGQGVASKIVGPSGAVPCK-----VEPG-----LGADNS 1029

OY

572

CV--LPDFLOGRATGRY-VGVYFDGLHPDSVSPFR-VAPLFSLEPTOLAPFIDLALOGGC 627

DB

1030

VYRFLP-----REGREYEVETTYDGVPIVPGS--PFLPLAVAN--TKPSKVKAFPGILOGG- 1080

OY

628

STSGARRPADREVERVTOALRSALDSCSSSEAPGCCCEMDLGPC 670

DB

1081

--SAGSPA-----RFTIDTKGAGCTGLGTF--GPC 1108

RESULT

9

STANDARD;

PRT;

922 AA.

W70T_MOUSE

Q9DZV7;

01-MAR-2002 (Rel. 41, Created)

01-MAR-2002 (Rel. 41, Last sequence update)

01-MAR-2002 (Rel. 41, Last annotation update)

70 kDa WD-repeat tumor rejection antigen homolog.

Mus musculus (Mouse).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxId=10090;

[1]

...

[illegible]

Query Match	Best Local Similarity	2.7%	Score 98.5	DB 1	Length 1447
Matches 97	Conservative 61	Mismatches 176	Indels 125	Gaps 21	
20 SLERLMEPDPTARCSLGLSCHLMDGVLCLEPGSLQSPAPVLPTRLRQTEVLV-----R 73					

or send an email to license@sib-sib.ch).

KA MEDLINE=95303480; PubMed=1784062;
RA Gray D.A., Inazawa J., Gupta K., Wong A., Ueda R., Takahashi T.;
RT "Elevated expression of unph, a proto-oncogene at 3p21.3, in human

Lung tumors. ;
Oncogene 10:2179-2183(1995).
[2]
REVIEWS.
Gray D.A.;
Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
MEDLINE=98124180; PubMed=9464533;
Frederick A., Rolfe M., Chiu M.I.;
The human UNP locus at 3p21.31 encodes two tissue-selective,
cytoplasmic isoforms with doubletinating activity that have reduced
expression in small cell lung carcinoma cell lines.;
Oncogene 16:153-165(1998).
-1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
ubiquitin + a thiol.
-1- SUBCELLULAR LOCATION: Cytoplasmic.
-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; UNPCL (SHOWN HERE) AND UNPES;
ARE PRODUCED BY ALTERNATIVE SPLICING.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19; ALSO KNOWN AS
FAMILY 2 OF UBIQUITIN CARBOXYL-TERMINAL HYDROLASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

EMBL: U20657; AAB72237.1; ;
EMBL: AF017305; AAC27355.1; ;
EMBL: AF017306; AAC27356.1; ;
MEROPS: C19.010; ;
MIM: 603486; ;
InterPro: IPR001394; UCH-2.
Pfam: PF00442; UCH-1; 1.
Pfam: PF00443; UCH-2; 1.
PROSITE: PS00972; UCH_2_1; 1.
PROSITE: PS00973; UCH_2_2; 1.
PROSITE: PS00973; UCH_2_3; 1.
Ubiquitin conjugation; Hydrolase; Thiol protease; Multigene family;
Proto-oncogene; Alternative splicing.
ACT_SITE 311 311
ACT_SITE 873 873
ACT_SITE 881 881
VARSELIC 232 279
BY SIMILARITY.
KSTRAPSRNFTTSPKSSASPSYSSASLIANGDSTSTGCMH
SSGVSRG -> N (IN ISOFORM UNPES).
C->A: LOSS OF ACTIVITY.
S -> R (IN REF. 3).
R -> S (IN REF. 3).
SEQUENCE 963 AA; 108564 MW; 1B62B752F9410CD7 CRC64;

Query Match 2.78; Score 97; DB 1; Length 963;
Best Local Similarity 20.18; Pred. No. 7.1;
Matches 107; Conservative 66; Mismatches 189; Indels 170; Gaps 26;

3 VSMFLSLALGRNVVSVLEKMEQDTRARCSLGSCHLMGDVCLPSGLQSPGVL- 61
133 VEYVLLLELKCEN-----SDPTNV-----LSCHFSKADPTI--AVTEKMRKLFN 174
62 VPRILQTELVLRCPKT-----DCAL--RVRVYVLAHVGHAPEPEACKSD 106
175 IPAREERLRNKNKTSNTYEQSLKNDNTVQDAGLYOGQVAVTEPQNDGTW--PROTLQSK 232
107 SELQESRN-----ASLQAVVLSFQAYPIARCALLEVQV----- 140
233 SSTAPSRNFTTSPKSSASPSYSSASLIANGDS--TSTGCMHSSGVSRSRGSGFSASYNCQ 290
141 --PADLVQPGQ-SVGSANVFQDFEASLGLAEVQWSTTKP-----RYQELNLTQOLPD 189
291 EPPSSHQPGGLGSLGNLGNMCF--MNSALQCLSTWAPLTLFDYFKDDEYEAEN----- 340

QY 190 GDNVL-LTLDVSEEDQDFSLTLRPVDAKSLM-----YKNLG---PONTTL 234
Db 341 RDNPLGKMGELAE-----AYAEILKQMSGRDAHVAHPSMFTQVGRFAPQPSG 389
QY 235 NHTDLVPCLCIQVMSLEPDSERVEPCF-----REDGARNHMLHIALRVLS--- 282
Db 390 QOQSOELLAEFLDLHDLHEDLVNKKPKYLELKDANGRPDAVVAKEAMENHRBDSYVD 449
QY 283 --PEWMQDAPCCLEPGKTYLCAQADQSPQCPPLVPYPQANAYNE----POD----- 329
Db 450 TEHGLFKSTLVCEPCAKSVTF-----DPCYVLTLLPLPKKDRMEVELVPADHCRPTQ 504
QY 330 ----FQVAGHPNLCVQVSTWEKVOLQACSMADSLG-PFKDMLVEMKGLNNT----- 379
Db 505 YRVVPLPLMGANSIDLCEALSRLSGIAENNVADVYHNRFRK---TFQMDCLNIMPDD 561
QY 380 ----SVC--ALEPSCGTPV-----PSMASTRARLGEELIDPFRSHQ 415
Db 562 IFVEYECSTVDGSECVTLPVYFRERKSRPSSSSALGQPLLVSVPKHK 613

RESULT 14
ID RYR3_HUMAN STANDARD: PRT: 4870 AA.
AC 015413; 015412; 015175;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ryanodine receptor 3 (brain-type ryanodine receptor) (RYR3) (RYR-3)
DE (Brain ryanodine receptor-calcium release channel).
GN RYR3 OR HBRR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND VARIANTS.
RC TISSUE-Petal brain;
RX MEDLINE=98175492; PubMed=9515741;
RA Leeb T., Brienly B.;
RT cDNA cloning and sequencing of the human ryanodine receptor type 3
RT (RYR3) reveals a novel alternative splice site in the RYR3 gene.;
RN FEBS Lett. 423:367-370(1998).
[2]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND TISSUE SPECIFICITY.
RC TISSUE-Brain;
RX MEDLINE=98055474; PubMed=9395096;
RA Nakashima Y., Nishimura S., Maeda A., Barsoumian E.L., Hakamata Y.,
RA Nakai J., Allen P.D., Imoto K., Kita T.;
RT Molecular cloning and characterization of a human brain ryanodine
RT receptor.;
RN FEBS Lett. 417:157-162(1997).
[3]
RP SEQUENCE OF 520-660 FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE-Skeletal muscle;
RX MEDLINE=98268728; PubMed=9607712;
RA Martin C., Chapman K.E., Seckl J.R., Aschley R.H.;
RT Partial cloning and differential expression of ryanodine
RT receptor/calcium-release channel genes in human tissues including the
RT hippocampus and cerebellum.;
RN Neuroscience 85:205-216(1998).
[4]
RP SEQUENCE OF 3943-4870 FROM N.A.
RC TISSUE-T-cell;
RX MEDLINE=95010709; PubMed=7523185;
RA Hakamata Y., Nishimura S., Nakai J., Nakashima Y., Kita T., Imoto K.;
RT Involvement of the brain type of ryanodine receptor in T-cell
RT proliferation.;
RN FEBS Lett. 352:206-210(1994).
[5]
RP SEQUENCE OF 4644-4842 FROM N.A.
RC TISSUE-Cervical carcinoma, and Hepatoma;

MEMLINE-94102751; PubMed-8276408; Sorrentino V., Giannini G., Malzac P., Mattei M.G.; "Localization of a novel ryanodine receptor gene (RYR3) to human chromosome 15q14-q15 by in situ hybridization."; Genomics 18:163-165(1993).

[6]

SEQUENCE OF 4652-4803 FROM N.A. TISSUE=Myometrium; MEMLINE-96032336; PubMed-7556644; Lynn S., Morgan J.M., Lamb H.R., Meissner G., Gillespie J.I.; "Isolation and partial cloning of ryanodine-sensitive Ca2+ release channel protein isoforms from human myometrial smooth muscle."; FEBS Lett. 372:6-12(1995).

-1 FUNCTION: Communication between transverse-tubules and sarcoplasmic reticulum. Contraction of skeletal muscle is triggered by release of calcium ions from SR following depolarization of T-tubules (By similarity).

-1 SUBUNIT: Homotetramer (Potential).

-1 SUBCELLULAR LOCATION: Integral membrane protein (Probable).

-1 ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; are produced by alternative splicing.

-1 TISSUE SPECIFICITY: Brain, skeletal muscle, placenta and possibly liver and kidney. In brain, highest levels are found in the cerebellum, hippocampus, caudate nucleus and amygdala, with lower levels in the corpus callosum, substantia nigra and thalamus.

-1 MISCELLANEOUS: The calcium release channel is modulated by calcium ions, magnesium ions, ATP and calmodulin.

-1 MISCELLANEOUS: The calcium release channel activity resides in the C-terminal region while the remaining part of the protein constitutes the 'foot' structure spanning the junctional gap between the SR and the T-tubule. It is possible that the foot structure interacts with the cytoplasmic region of the dihydropyridine receptor.

-1 MISCELLANEOUS: Ryanodine is an alkaloid that binds to the Ca-release channel in junctional SR and modulates its activity.

-1 SIMILARITY: BELONGS TO THE RYANODINE RECEPTOR FAMILY.

-1 SIMILARITY: CONTAINS 3 SPRY DOMAINS.

-1 CAUTION: Ref.2 sequence differs from that shown due to probable frameshift errors at position 742-766.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AJ001515; CA04798.1; -
EMBL; AB001025; BAA23795.1; -
EMBL; AJ002512; CA05503.1; -
EMBL; X74269; CAA52326.1; -
EMBL; X74270; CAA52327.1; -
MTM; 180903; -
InterPro: IPR000636; Cation_chan_non_lig.
InterPro: IPR001682; Channel_pore_Ca_Na.
InterPro: IPR002048; EF-hand.
InterPro: IPR003608; MIR.
InterPro: IPR000699; RYDR_ITPR.
InterPro: IPR003032; RYR.
InterPro: IPR001215; Ryanodn_receptor.
InterPro: IPR003877; SPRY.
InterPro: IPR003878; SPRY_domain.
Pfam; PF00036; efhand; 1.
Pfam; PF00520; ion_trans; 1.
Pfam; PF02815; MIR; 4.
Pfam; PF01365; RYDR_ITPR; 2.
Pfam; PF02026; RYR; 4.
Pfam; PF00622; SPRY; 3.
PRINTS; PR00795; RYANODINER.
SMART; SM00472; MIR; 4.
SMART; SM00449; SPRY; 3.
Receptor; Transmembrane; Ionic channel; Calcium channel; Repeat;

Query Match	Best Local Similarity	Conservative	Score 97	DB 1	Length 4870
QY	6	FLTLSTLGRNPVVSLERLMEPODPAKSLGSGC-TM-----DGDVLCPGSLQSAFG	58		
Db	957	YVMSNNTKRPADLDSDVTKLPPELIVYKLAHNAHVMYAKDRKQGTYYGIQDPLKRN	1016		
Y	59	PVLPYRLQTELVRCPDKQIDCALRVRRVYVHLAHGMAPEEAGKSDSELSRNSAQ	118		


```

Db 1017 PRLVYVYALDE---RTKKSNRDSLREAVRTFVG-YGYNIER-----SDQELADSAVERKVS 1067
OY 119 AQVVLSE---QAVPIARCAL---LEVOVPADLVOPGQSVGSVAFDCE--EASIGAEVOIMS 171
Db 1068 IDKIRFERVERRSYAVNSGMYEFEEVYTGDM-----RVGNARPCRDVEYEGADQAFV 1122
OY 172 YTKPRYOK-----ELNLQQLPDGDNVLLTLDVSEODFSFLLYLRVPVPAKLSMYKN 225
Db 1123 FEGNNGORHOGSGYFGRTWQPDGVVCGMINLD-----DASMIFF-----T 1162
OY 226 LTGPQNTLNTHTDL-----VPCLCIOV-----WSLEPDSERVEF--C----- 260
Db 1163 LNGELLITTKGSELAFADYEIENGFPVICIGLSQIRMINIGTASTKFTYMCGLQEGF 1222
OY 261 -PFRDPCGHRNLM-----HIAIRVLSPCWQOLDAPCCLPKTYT----- 299
Db 1223 EPFAVNMNRDVAMVMSKRLPTFVNVPRKDHIEVMRIDG-----TMDSPCL--KYTHKTF 1276
OY 300 -----LCWQAPDQSPCQPLVPPVPQKNATVNEPQDFOLVAGHPNLCV 341
Db 1277 GTQNSNADMIYCRLSMPVECHSFSHSPC--LDSEAFQKRQME-----ILSHTTQCY 1329
OY 342 QVSTWEKVQIACSNADSLGPRFDDMLVEMKTGLNNTSCALBPSGCTPLPSMASTRAA 401
Db 1330 YAIRIFAGQDPSCVWVGWVTP--DYHLXSEKEDLN-----KNCTVTVTLDDER-G 1376
OY 402 RIGELLQDFRSHQCMQNLNDDNMGSLMACPMKXIIHRWVLYVLACLL-LAAALFFELL 460
Db 1377 RVHESV-----KSNCTYMWGGLVVS-----SQRNSRNVDELICGLVDLAMEGLSF-- 1424
OY 461 LKQRRKARQSR-----TALLHSADAGYERIVGALASALSOMPLRAYV 506
Db 1425 -----SANGKELCTCYOYENPTKYFPAFLQPTSTLSLOFELGKLKNA--MPLSAI 1474
OY 507 DLMSRRELHANGALMFIHQRRIIDEGGVILLFSPAAYACQOMLOLQVPER 560
Db 1475 -----FSEKKN-----PVQCPRLDVOYIOP 1497
RESULT 15
Y2ML MOUSE STANDARD: PRT: 1476 AA.
P28665;
01-DEC-1992 (Rel. 24, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Moringlobulin 1 precursor (MUG1).
MUG1 OR MUG-1.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A., AND SEQUENCE OF 28-57.
TISSUE=Liver;
MEDLINE=91358495; PubMed=1840592;
Overbergh L., Torrekens S., van Leuven F., van den Bergh H.;
Molecular characterization of the murinoglobulins.;
J. Biol. Chem. 266:16903-16910(1991).
J. FUNCTION: A PROTEINASE ACTIVATES THE INHIBITOR BY SPECIFIC
PROTEOLYSIS IN THE BAIT REGION, WHICH, BY AN UNKNOWN MECHANISM
LEADS TO REACTION AT THE CYSTEINYL-GLUTAMYL INTERNAL THIOL ESTER
SITE AND TO A CONFORMATIONAL CHANGE, WHEREBY THE PROTEINASE IS
TRAPPED AND/OR COVALENTLY BOUND TO THE INHIBITOR. WHILE IN THE
TETRAMERIC PROTEINASE INHIBITORS STERIC INHIBITION IS SUFFICIENTLY
STRONG, MONOMERIC FORMS NEED A COVALENT LINKAGE BETWEEN THE
ACTIVATED GLUTAMYL RESIDUE OF THE ORIGINAL THIOL ESTER AND A
TERMINAL AMINO GROUP OF A LYSINE OR ANOTHER NUCLEOPHILIC GROUP ON
THE PROTEINASE, FOR INHIBITION TO BE EFFECTIVE.
SUBUNIT: MONOMER.
TISSUE SPECIFICITY: PLASMA.
SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M65736; AAA73048.1;
DR PIR; A41185; A41185.
DR HSSP; P01023; 1BV8.
DR MCD; MGI:99837; MUG1.
DR InterPro; IPR002890; A2M_N.
DR InterPro; IPR001599; Alpha_2_macroglbln.
DR Pfam; PF00207; A2M_N.
DR Pfam; PF01835; A2M_N.1.
DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
KW Serine protease inhibitor; glycoprotein; plasma; bait region; signal;
KW Multigene family.
FT SIGNAL 1 27
FT CHAIN 28 1476
FT DOMAIN 677 734
FT DISULFID 48 86
FT DISULFID 251 276
FT DISULFID 269 288
FT DISULFID 461 555
FT DISULFID 587 773
FT DISULFID 634 680
FT DISULFID 849 885
FT DISULFID 923 1323
FT DISULFID 1081 1129
FT DISULFID 1354 1469
FT THIOLEST 974 977
FT CARBOHYD 55 55
FT CARBOHYD 294 294
FT CARBOHYD 313 313
FT CARBOHYD 500 500
FT CARBOHYD 749 749
FT CARBOHYD 776 776
FT CARBOHYD 871 871
FT CARBOHYD 993 993
FT CARBOHYD 1142 1142
FT CARBOHYD 1180 1180
FT CARBOHYD 1426 1426
SQ SEQUENCE 1476 AA; 165139 MW; D145D325B2DDDEAB CRC64;
Query Match 2.7%; Score 96.5; DB 1; Length 1476;
Best Local Similarity 20.5%; Pred. No. 14;
Matches 76; Conservative 47; Mismatches 161; Indels 87; Gaps 17;
OY 67 QTELVLRCPO-----KTDCALRVRVYVLA-----HGHWAEPEEA--- 102
Db 454 ETSSILPCNQIHTVQAHFILKGDGLVLRKELIFVLYMAGSITOTGHTHOVEGEAPVK 513
OY 103 GKSSSELQESRNASLQAOVLSFOAVPIARCALLEVOVPADLVOPGQSVGSVAFDCEAS 162
Db 514 GK-----FALRIPVEFSWVPKMKILYTLIPDQEV--ADVNFRIKCLRNK 559
OY 163 LGAEVQIWSYTKPRQKELNLQQLPDGDNVLLTLDVSEODFSFLLYLRVPVPAKLSM 222
Db 560 VDLRFST--SOSLPASQTRLOVTA--POSICGLRAYDOS-----VLLLRPESE-LEPSW 609
OY 223 YKNLTGPQNTLNTHTDLV-----CLCIOV-----SLEPDSERVEFCPFREDP 266
Db 610 IYNLPQGQ-----QNKFPVSSRLSDQEDCIIYSSSLAKHNLVHGCTKDYRYVEDM 664
OY 267 G--AHRNLMHIALRVLSPCWQOLDAPCC--LPKVTLCWQAPDQSPCQPLVPPVPQKN 321
Db 665 GLTAFTNIMIKRLPIYCFDYGWVPISAPRVFPIAFTPEISWSLRTLSKRPREP--PRKD 722
OY 322 ATVNEPQDFQLVAGHPNLCVQVSTWEKVQIACSNM-----DSLGRFDDMLVEMKT 374

```


Db	723	PSSNDPLFETIRKPEPEWV---	NDIVVNSGIAEVENTVPDITLWKAGALCISNDT	778
Oy	375	GLNNTSYCALE	385	
Db	779	GLGSSVPLQ	789	

Search completed: September 28, 2002, 19:43:34
Job time: 282 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

3M protein - protein search, using sw model

Run on: September 28, 2002, 17:59:47 ; Search time 43.34 Seconds
(Without alignments)
1494.329 Million cell updates/sec

Title: US-09-899-471-2
RefSeq score: 3605
Sequence: 1 MPVSWFLSLALGRNPVVS.....SSEAPGCCBEMDLGPTTLE 674

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR_71: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	3.2	267	2 A38442	probable tumor sup
2	114.5	3.2	1013	2 B70841	probable helz prot
3	110	3.1	2910	2 T42214	otogelin - mouse
4	108.5	3.0	361	2 T39784	hypothetical prote
5	106.5	3.0	478	2 D75564	hypothetical prote
6	105.5	2.9	698	2 A62593	hypothetical prote
7	105.5	2.9	1451	2 B41185	alpha-2 macroglobu
8	105	2.9	1615	2 BA9502	protein-tyrosine-p
9	105	2.9	1767	2 A49502	protein-tyrosine-p
10	103.5	2.9	3707	2 S18252	heparan sulfate pr
11	103	2.9	851	2 T31520	hypothetical prote
12	103	2.9	906	2 G83156	probable transcrip
13	102.5	2.8	1711	1 A55148	protein-tyrosine-p
14	101.5	2.8	1476	2 A41185	alpha-2 macroglobu
15	100	2.8	1447	2 A54100	tumor suppressor p
16	99.5	2.8	2647	2 A37098	gelation factor AB
17	99	2.7	1617	2 T28153	complement C4 - ch
18	99	2.7	2437	2 S42612	transmembrane prot
19	99	2.7	3014	1 JC5620	genome polyprotein
20	98.5	2.7	665	2 S69222	probable transcrip
21	98	2.7	824	2 H90647	ATP-dependent heli
22	98	2.7	824	2 H85498	helicase, ATP-depe
23	97.5	2.7	1618	2 S21424	nestin - human
24	97.5	2.7	1763	2 T17465	rifamycin polyketi
25	97.5	2.7	1894	2 C54689	protein-tyrosine-p
26	97.5	2.7	2629	2 T30987	telomerase-associa
27	97	2.7	604	2 G83091	conserved hypochet
28	97	2.7	963	2 T09478	ubiquitin thiolest
29	96.5	2.7	813	2 AF0526	ATP-dependent heli

30	96.5	2.7	1187	2 JF0347	hypothetical prote
31	96.5	2.7	1691	1 D54689	protein-tyrosine-p
32	96.5	2.7	2535	2 T04824	hypothetical prote
33	96	2.7	564	2 D96017	probable peptidase
34	95	2.6	1573	2 S01845	DNA (cytosine-5')-
35	95	2.6	10223	2 T30225	polyketide synthas
36	94.5	2.6	445	2 E70854	probable membrane
37	94.5	2.6	837	2 A34898	granulocyte colony
38	94.5	2.6	6260	2 T30228	polyketide synthas
39	94	2.6	509	2 JC5288	SHP substrate-1 pr
40	94	2.6	513	2 JC5289	SHP substrate-1 pr
41	93.5	2.6	927	2 A48085	transcription fact
42	93.5	2.6	1513	2 T23681	hypothetical prote
43	93.5	2.6	6831	2 A88852	protein unc-22 (lm
44	93.5	2.6	6839	2 S57242	twitchin (stimlari
45	93.5	2.6	7160	2 T27935	hypothetical prote

ALIGNMENTS

RESULT 1
A38442
probable tumor suppressor - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 18-Oct-1991 #sequence_revision 18-Oct-1991 #text_change 16-Feb-1997
C:Accession: A38442
R:Nigro, J.M.; Cho, K.R.; Fearon, E.R.; Kern, S.E.; Ruppert, J.M.; Oliner, J.D.; Kinz
Cell 64, 607-613, 1991
A:Title: Scrambled exons.
A:Reference number: A38442; MUID:91121517
A:Accession: A38442
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-267 <NIG>
A:Cross-references: GB:M63696; GB:M63700; GB:M63702; GB:M63718; GB:M63698
C:Keywords: transmembrane protein

Query Match 3.2%; Score 116; DB 2; Length 267;
Best Local Similarity 25.6%; Pred. No. 0.13;

Matches 72; Conservative 31; Mismatches 100; Indels 78; Gaps 16;

QY 155 VFDCFEASLGAEOVWSTYKRYQKELNLTQQLPD-----GDNVLLTLDVSEEDDFSL 209
DB 8 LYOC-EASLGGSGSIISRTAKVAVAPLRLSOTESYTAFGDVLKCEVIGE----- 59
QY 210 YLRVPPALAKSLWKYNLGPONITLHTD-----LVPCCLQVWSLSPDSERVEFCPPRED 265
DB 60 ---PMPRT---HMQRN---QODLPPIGDSRVVLPSPGALQISRLDGDIGIRCSAR-N 109
QY 266 PGARNLMLHRLRYLSPGVWOLDAPCCIPGKVTLCWOAPD-OSPCOPLVPPV-----OK 320
DB 110 PASRT-GENEVEVRLSP-PFELNHPNL-----YKESNDIEFECTVSGKPPVPTVMKK 162
QY 321 NATVNEPQDFOLVAGHPNLCAOVSTWEKVOLOACSNADSLGPKDD---MLLVEMKTGL 376
DB 163 NGDVVPSDFQIVGGSMRLI-----LGVVKSDEGFQVCAVENEAGN 204
QY 377 NNTSVCLPSCGTPPLPS-----MASTRARL 403
DB 205 AOTSALIVK---PIESSVLPSAPRDVVPVVLVSSRFVRL 242

RESULT 2

B70841
probable helz protein - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: B70841
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Steeger, K.; Skellton, S.; Squares, S.

Figure 393, 537-544, 1998
 Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 Reference number: A70500; MUID:98295987
 Accession: 870841

Status: preliminary; nucleic acid sequence not shown; translation not shown

Molecule type: DNA

Accession: 1-1013 <COL>
 References: GB:AL021924; GB:AL124346; NID:g3261519; PIDN:CA117284.1; PID:el25237

Experimental source: strain H37RV
 Genetics: helz

Query Match 3.2%; Score 114.5; DB 2; Length 1013;
 Best Local Similarity 20.4%; Pred. No. 0.91;
 Matches 143; Conservative 75; Mismatches 232; Indels 251; Gaps 33;

```

90 LAVHGHNAEP--EAGKSDSELOSERNASLQAVVLSFGAYPIAKCALLEVOPADL---144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2 LVHGFMSNSGGMRLMADSDDL---VKSPGQALSARHPFA-----APADLIAG 49
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
145 VORQSGSAVDFCEASLGAEVQIMSTKPKRYQKELNLTQQLPDGDNVLLITLVSEED 204
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
50 IHPRK-----PATAVLLPSLSAPLD 71
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
205 FSPFLYLRVPVDALKSLMYKNLTGPONITLNTD-LVPCLCIQVMSLEPPSERVEPCFR 263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
72 SPFLIRLAPRAA-----RDPMLAMTVVVDLDPDAALAAFPDPA 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
264 EDCGHNHMLHIALRLVLSPGWQOLDAPCCLPCKY-----TLG-----QAPDQS 308
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
114 PDVRYGASVDLAEIAYARELVE-----KGRVLPOLRRDTGGAACRPVLYQGRVY 166
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
309 PCOPLY---PVPQKNATVNEPQDFOLVA-----GHPNLCVQ 342
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167 AMTSLVSAHPVCAEVGHPHEIATSLDAMVDAVRAALSPMDLPPRRGSKHRA 226
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
343 VSTW-----EKVQLQACS-----WAD-----SLGPKKDMILVEMKTGLNNTS 380
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
227 VEAMLTALTCPPDRFPDAPDELDALEALRPMDVVGIGTVGPARATFRLESEVENTEET 285
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
381 VCALPESCTPLPMASTRARARLSEELLQDFRSHQCM---QLMNDNMGSLMACPMDKY 436
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
286 -----PAG-----SLWRL-EFLQSTQDPSILVPAEQAMND--GSL----- 319
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
437 IHRKWV---LWMLACLILAAALFFLLKKDRKRAKGSRTALLSHSADAGAYERLVGA 492
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
320--RRWIDRPOELLITELGRASRIPELV-----PALRTACPSGLEIDADCA--YRFLSG 369
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
493 LASALSGMPLRVAVDWMSRRRLSAGALAMFHHORRRIQEGGVILLSESP--AAVAQCQ 550
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
370 TAAVLDEAGEGVLLPSW-----W-----DRRRKL--GLVLSAYTPVDGVGKAS 411
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
551 QMLQQLQVE-----PGPHDALAAMLSCLP-----DPLQG 580
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
412 KFGREQLVFEFWEIAGVDDPLSEEEIALITERKPSILRLRGQVVALDTEQMRGLEFLER 471
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
581 RATGRVYGVYFDGLL--HPDSVPSPFRVAPLPSLTQLPALFDALQGGCSTSGRPADEV 638
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
472 KPRGRKTTAETIALAASHPDVDVPLEVTAV-----RADGMGLDILLAGAAAASIQPDPP 526
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
639 ERVYQALRSALDSTS-----SSEAPGCGCEEMDLGPTTLE 674
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
527 DGTATILRPYQGRGLAMLAFLSSIGLSCGLADMGIGKTVQ 567
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 3
 10214
 10214
 Intername names: mucin-like extracellular matrix protein
 Species: Mus musculus (house mouse)
 Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: T42214
 R:Cohen-Salmon, M.; El-Amraoui, A.; Leibovici, M.; Petit, C.
 Proc. Natl. Acad. Sci. U.S.A. 94, 14450-14455, 1997
 A:Title: Otogelin: A glycoprotein specific to the acellular membranes of the inner ear
 A:Reference number: 222079; MUID:98070772
 A:Accession: T42214

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2910 <CON>

A:Cross-references: EMBL:U96411; NID:g27260883; PID:g27260884; PIDN:AAB96561.1

A:Experimental source: strain BALB/c

A>Note: component of all the acellular membranes of the inner ear

C:Superfamily: Von Willebrand factor; von Willebrand factor type A repeat homology; v

Query Match 3.1%; Score 110; DB 2; Length 2910;
 Best Local Similarity 20.8%; Pred. No. 7.9;
 Matches 54; Conservative 31; Mismatches 82; Indels 92; Gaps 13;

```

OY 214 VPDLKSLMYNNINGPQNTLNTHTDLVPCLCIQVMSLEPPS-----ERYEPCFR 263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2427 VPRALGETWNSLSG-----CCQOCCQAPDTIIPVDLDCRGPRESGP-- 2469
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 264 EDCGHNHMLHIALRLVLSPGWQOLDAPCCLPCKYTLCMQAPDQSPQPLVPVQKNAT 323
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2470 -----RGEVYILQP-----TEDPCL-GSVCVC-----NQTLCGLAPTCRPGHSL 2510
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 324 VNEPQDFOLVAGH-----PNLCQVSTWEKVOLOACSNADSLGPFKDMILVEMKTGLN 378
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2511 ITHQEDSCSPSYSCQCPGLC-----EAEVPTC-----RNDQLILBERLG-- 2552
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 379 TSVCALEPSCG---TPLPMASTRA-----ARIGEELQDFRSHQ----- 415
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2553 DSCCTSTFCGCGSDPBEQCEGELTVHNTTELCPLYQCVCENFRCPVOVCGMGT 2612
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 416 CMQMDNDNMGSLMACPMD 434
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2613 LVEYWSPPDCCPKYKSCD 2631
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 4
 T39784
 hypothetical protein SPBC18H10.20C - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T39784
 R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Badcock, K.; Churcher, C.M.
 submitted to the EMBL Data Library, March 1998
 A:Reference number: Z21879
 A:Accession: T39784
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-361 <LYN>
 A:Cross-references: EMBL:AL022304; PIDN:CA118417.1; GSPDB:GN00067; SPDB:SPBC18H10.20C
 A:Experimental source: strain 972h; cosmid c18H10
 C:Genetics:
 A:Gene: SPDB:SPBC18H10.20C
 A:Map position: 2

Query Match 3.0%; Score 108.5; DB 2; Length 361;
 Best Local Similarity 20.3%; Pred. No. 0.74;
 Matches 84; Conservative 53; Mismatches 133; Indels 143; Gaps 21;

```

OY 16 PUVVSLERLMEPODPARCSLGLSCHLMGDVYCLPGSLQSAFG-----PVL 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 PLKIALPRSTTPKDPACTIDIR---MESPLVFLGSPETSSGALASGLIKLITLHOPFI 58
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 62 VPRILQELVLRCPQKTDCAIARVVVHLAVGHWAPEEAGKSDSELOSERNASLQAV 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 59 KVTTLKLIQILIKRI-----TLHLPAL-SHCS-----ACAGSKREVQTMPLANTTY 102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 122 VLSFGVPIAKCALLEVOPADLVQPGSGVSAVDFCEASLGAEVQIMSTKPKRYQKEL 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


[illegible]

RESULT 8
 B49502
 protein-tyrosine-phosphatase (EC 3.1.3.48) receptor type 4E, splice form B precursor
 C:Species: Drosophila melanogaster
 C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 24-Apr-1998
 C:Accession: B49502
 R:Oon, S.H.; Hong, A.; Yang, X.; Chia, W.
 J. Biol. Chem. 268, 23964-23971, 1993
 A:Title: Alternative splicing in a novel tyrosine phosphatase gene (DPT4E) of Drosophila
 A:Reference number: A49502; M0ID:94043220
 A:Accession: B49502
 A:Molecule type: mRNA
 A:Residues: 1-1615 <COND>
 A:Cross-references: GB:L20894
 C:Genetics:
 A:Gene: FlyBase: Ptp4E
 A:Cross-references: FlyBase: FBgn0004368
 A:Introns: 1605/3
 C:Superfamily: protein-tyrosine-phosphatase, receptor type 4E; fibronectin type III r
 C:Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; rec
 F:1234-1270/Domain: transmembrane #status predicted <TMN>
 F:1271-1615/Domain: intracellular #status predicted <INT>
 F:1353-1573/Domain: protein-tyrosine-phosphatase homology <PtpI>
 F:1525/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:1531/Binding site: substrate phosphate (Arg) #status predicted

Query Match 2.9%; Score 105; DB 2; Length 1615;
 Best Local Similarity 22.3%; Pred. No. 9.4;
 Matches 98; Conservative 49; Mismatches 156; Indels 156; Gaps 26;

QY	60	VLVPTRLQLELVLR	-----	PKTKDCAIRRVVHL	-----	AVHGHWAEPEEA	102
DB	403	ILEPGTLY	EVVYKTTADNVN	SWPASEVTILRPPRVRS	LAGFLDRSNALHISW	EPAEY	460
QY	103	GKSDS	-----	ELQESRNMS	-----	LOAOVVLSEQAVPI	145
DB	461	GKDSIRISYHEQTS	EVAVAPPVAAESQIT	TMLTETLDSLAGR	RYLNAVQALS	---	517
QY	146	QPGQSGSAVFDFC	EASLGAENVQIMSYTKR	YQKELNLTQOLP	DDNVLL	---	202
DB	518	---KGVASNASD	---	ITRYTRP	---	AAPLIQELRSIDQ	GLMSRSDVNSR 559
QY	203	QD	-----	FSEFLYLRP	-----	VPDAIKSLMYKNLTG	228
DB	560	QDRYEVHYQNRG	TRERTMATNETSLTI	HYLHPDSGYEVKVA	HAISHGRVSRSEPSYQAVF		619
QY	229	---PONITLNL	---HTDVLPPCLCIQW	SLSEPDSE	---	RVEECPFRPDPAHNR	271
DB	620	PKRPONLTQIQT	YNTNLV	---	VLHMQAPSGSDSE	EYVYRYRTDAS	PMQRISLHENEARI 675
QY	272	---	LMIHARLRLVSLPGV	WQMDAPCCLP	CKATLQWQAPDOS	PCQPLVPVPQKNATVN	325
DB	676	KDMHYGERLYLVQ	NVTFSGV	---	ESHPRLNLVY	---	PPQ
QY	326	EPQDDQLVAGHP	RLCVQVSWKEVQ	LOACSHADSLG	PFKDKMLVLEAK	---	IGLANTSVCA 383
DB	728	WFRP	---	DGH	---	VDFTYTLKMWPT	---
QY	384	LEPGCTPLPSMA	STRAR	402			
DB	764	SPSVRIPIEDLS	SPGRQYR	781			

RESULT 9
 A49502
 protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type 4E, splice form A precursor
 C:Species: Drosophila melanogaster
 C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 24-Apr-1998

molecule type: mRNA

Db 663 QLSEE 667

81174-1398/Domain: protein-tyrosine-phosphatase homology <PP1>
81350/Active site: Cys (phosphotyrosine intermediate) #status predicted
81356/Binding site: substrate phosphate (Arg) #status predicted

Query Match 2.8%; Score 102.5; DB 1; Length 1711;
Best Local Similarity 20.6%; Pred. No. 16;

Matches 149; Conservative 79; Mismatches 229; Indels 267; Gaps 40;

```

43 DGDVLCFGSLDQA-----PGVLPVTRIQTEL-VLRCPOR-----TDC 80
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
333 DGVYKLGSPMESTIGDECAVAVPPG-LRPGHYTLQKVLGAPYAWEGSTWLMS 391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
81 ALRVVYVHVLAVHGMAREEAGKSDSELSQESRNASLQAOVYLSFOAYPIARCALLEVOY 140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
392 AALPVEVGAFL--WLDGLEASK-----QPGRRALLYSQ--DAPSGISGNISV 434
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
141 PADLVQ-----PGQ-----SVGSAVECFEASLGAEOVQMSYTKPRQKELNLQQLPD 189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
435 PSGAHTVIFCGLVPGAHYKVDIASSTGDISQISG-----YTSPLPQSLV----- 481
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
190 GDNVLLTLVSEEDQFSLLYLRVPDAL--KSLWYKNLT-----GPQNTILNH 236
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
482 -----ISRSSPSDLTIANGPAPGQLEGYKVTMHQDSQSRPGDLVDGPDLSITL 532
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
237 TDLYP--CLCIQVW--SLEPDSERVECP-----FREDPGAHNLNH----- 274
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
533 KSLVPGSSYTVSAMAMAGMIGSDQKIHSCTRAPPTNLSLGFANPAAKASVHYHPGG 592
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
275 -----IARLRVL--SPGY-----W-QLDAPCCPGKATLWQADQSPCQPLV 314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
593 RDAFHRLRLRLPLTLESEKVPREAKQNFSAQLTAGCEFOVQDLSLTMSGRSS----- 646
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
315 PPVQKATVNEPQDFOLVAGHPNLCVOYSTWEKVOLOACSNADSLGPFQKDMILVEKMT 374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
647 -----SANAATGTP-----PSAPTL-VNYSQAPYQLOV-SMAHVPG----- 681
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
375 GLNNTSYCALPESGCTPLPSMASTRARLGEELLQDFRSHOQMLMNDNGSLMACPD 434
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
682 GRSRYQVTLVQESTRTATSIMGP-----KEDGSFGLTGT 718
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
435 KYIHRRWLVLAACLLAA--LEFFILLKKDRKRAKRSRTALLHSHADGATERLV- 490
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
719 KY--KVEYISAGPLTYTAANVSANTYPLIPNE-----LIVSMQASAVVNLAW 765
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
491 -----GALASALS-----QMPLRVAVDLMSRRELSHAGALAMFHRRRILOEGV 536
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
766 PSGLGCGACAHQSLDAGHLSWEOPRLKQGLFRLMDLTPGHITSMVKRCRAGFLQASTH 825
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
537 VILFSAVAQAQCOQMLQLOVPEGP-HDAL-----AAMLS-----CYLPPELQ 579
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
826 LVVL-----SVPEPVEDVLCHEPXYTLALNTMPAGDVYCLV----- 864
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
580 GRATGRVY-----GYE-----DGLHPDVSF--SPFRVAPLPSLPLQALFALDAGCC 627
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
865 --VVERLYVPGGCTHEFOVNTSGDALLPLNLPMTYSRLS--LTVLGRNSRWSRAVSLVC 920
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
628 STSA 631
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
921 STSA 924
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1476 <OVER>
A:Cross-references: GB:M65736
C:Superfamily: alpha-2-macroglobulin

Query Match 2.8%; Score 101.5; DB 2; Length 1476;
Best Local Similarity 21.4%; Pred. No. 16;

Matches 72; Conservative 44; Mismatches 149; Indels 71; Gaps 16;

```

86 VVHILAV-----HGMAREEAA--GKSDSELSQESRNASLQAOVYLSFOAYPIARCALLE 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
489 VMAHLSITQGNHTHQVEGEAPVKK-----PALEIPVRSVPMKMLYT 536
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
138 VQVPADVLQPOGQSGSAVDFCEASLGAEOVQMSYTKPRQKELNLQQLPDGDNVLLTL 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
537 ILPDGEVTL--ADSVNPEIEKCLRNKVDLRFST--SGLPASQTRQVYAS--PQSLGRLAV 592
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
198 DVSEEDQFSLLYLRVPDALSLWYKNLTGPQNTILNHTDLYP-----CLCIQV 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
593 DOS-----VLLKPESE--LSPSWIYNLPQM-----QMKFVPSRSLSEDOEDCILYSS 639
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
248 W-----SLEPDSERVECPFREDPG--AHRNLMLHARLRLVSPGWLDPAPCC--LPG 296
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
640 WLAERHNLVPHGTEKDYRYVEDMGLTAFNTLMIKLPIICFDGWPISAPRVEDLAF 699
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
297 KYTLCNQAPDQSPCQPLVPVQKATVNEPQDFOLVAGHPNLCVOYSTWEKVOLOACSW 356
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
700 TPEISMSLRTLLSKRPEEP--PRKDPSSNDPLRETIIRKYPEETWV--MDIVYVNSTGL 753
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
357 A-----DSLCPFKDMLVEMKGTGLNNTSVCALE 385
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
754 AEVEMTVPDITTEMKAGALCLSNDTGLSSVPLQ 789
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15

tumor suppressor protein DCC precursor - human
N:Alternate names: colorectal cancer suppressor DCC
C:Species: Homo sapiens (man)
C:Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 05-Nov-1999
C:Accession: A54100; A40098
R:Hedrick, L.; Cho, K.R.; Fearon, E.R.; Wu, T.C.; Kinzler, K.W.; Vogelstein, B.
Genes Dev. 8, 1174-1183, 1994
A:Title: The DCC gene product in cellular differentiation and colorectal tumorigenesis
A:Reference number: A54100; MUID:95011532
A:Accession: A54100
A:Molecule type: mRNA
A:Residues: 1-1447 <HED>
A:Cross-references: EMBL:X76132; NID:9453209; PIDN:CAA53735.1; PID:9453210
R:Pearson, E.R.; Cho, K.R.; Nigro, J.M.; Kern, S.E.; Simons, J.W.; Ruppert, J.M.; Hami
Science 247, 49-56, 1990
A:Title: Identification of a chromosome 18q gene that is altered in colorectal cancer
A:Reference number: A40098; MUID:90100559
A:Accession: A40098
A:Molecule type: mRNA
A:Residues: 1-750 <FEA>
A:Cross-references: GB:M32922; NID:9181492; PIDN:AAA35751.1; PID:9181493
C:Genetics:
A:Gene: GDB:DCC
A:Map position: 18q21.1-18q21.1
C:Keywords: transmembrane protein; tumor suppressor
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-1447/Product: tumor suppressor protein DCC #status predicted <MAN>

Query Match 2.8%; Score 100; DB 2; Length 1447;

Best Local Similarity 18.2%; Pred. No. 21;

Matches 125; Conservative 56; Mismatches 183; Indels 322; Gaps 29;

70 LVLRCPQKTCALRV-----RVVVLAVHGMAREEAGKSDSELSQESRNASLQAOVYLSF 125

Dh	57	VLDDSAEDRCGPVIAKMKDKGJHIAL-----	GMDERKQQLNSGSLILIONILHS	105
QY	126	CAFPYARCALLEVQPADVLVOPQOSGSAVFDEEASIGAEVQWISYTP-KREYQREKMLT	184	
Dh	106	RHH-----KPRDE-----GLYQC-EASIGDSGSIISPTAKVAVAGPRLRL	143	
QY	185	QOLPD-----GDNVLLITLDVSEBODPSFLYLKRPVPOALKSLMYTKNLTPGONITLNMHD-	238	
Dh	144	STESVTAFFMGPVLVILKCEYIGE-----PMPTI-HWOKN---OODLTPTIPGDS	187	
QY	239	--LVPCLCIGWMSLEPDESERVEFCPPR-----EDGAGRHLMIH----	275	
Dh	188	RVVYLPFSGALQISRLQPGDIGITRCSARNAPASSKTGNEAEVRIISDPBMLHOLYFLQPS	247	
QY	276	-----	275	
Dh	248	NVAIEGKDAVLCECVSGYPPPSFTMLRGEBVIQLRSKKEYSLGGSNLLISNVTTDDSGM	307	
QY	276	-----ARKVLSGCVQOLADPCLPGKYVTLCTWQAPD-OSPCOPLVRVP-	318	
Dh	308	YTCVVYTKNENISASSELVYLVPE-MPFLNHPNSNL-----YAESMDIEFECTVSQGRVPT	361	
QY	319	---OKNAFVNEHPODFOLVAGHPMLCYQVSVWEKQVLQACMSADSLSGPEPKD---MLVE	371	
Dh	362	VNMKNGDVYIPSDYFQIYGGSNLRI-----LGVYKSDBEGYQVAAE	403	
QY	372	MKTGLNNTSVCALEBPSGCTP-----LPMSASTRAARLGEELLDFDRSHQCMQ.	419	
Dh	404	NEAGNAQTSQAQILVPRPALIPSSSVLPSAPRDVYVPLVSSRVRLS-	448	
QY	420	WNDDMGMSIACPMOKYTHRRWVLMVLAELLALLAALFFLLILKDKDRKAAGS-RTALLL	478	
Dh	449	-----WRPPAE-----ARGNIOTFTVE	465	
QY	479	HSADGAGYERVLGALASALSOMPLR-----VAVDLMSRRELISAGALAMFHQ	526	
Dh	466	FSRBEDNNERALTNPQPSGLQTLGVNLKPEAMTTFRVAVATINEMGPGESS-----	514	
QY	527	RRRILOEGGVILLFSPAVAOCQOMQO-----LQTVBERPHDALAMVLSVLPDFLO	579	
Dh	515	-----QPIKVA--TQPELQVGPVENIQAVSPTSILITWBP--PAVAN	555	
QY	580	GR-----ATGRYGVYFEDGL	594	
Dh	556	GPVQGRILPCTEVSTGKEONIEVDGL	581	

```

search completed: September 28, 2002, 19:41:27
job time: 6100 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2002, 19:39:44 : Search time 62.3 Seconds

(without alignments)
1244.454 Million cell updates/sec

Title: US-09-899-471.5
Perfect score: 3741
Sequence: 1 MPVSWFLSLALGRNPVVS.....SSEAPCCCEMDLPCTTLE 698

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2389.5	63.9	692	AAB61880	Human cytokine rec
2	2376	63.5	705	AAU029322	Human PRO polypept
3	2376	63.5	705	AAU04956	Human Interleukin
4	2376	63.5	705	AAB87606	Human PRO20040. H
5	2376	63.5	705	AAB61884	Chimeric Zcytor14
6	2314	61.9	675	AAB61885	Chimeric Zcytor14
7	2300.5	61.5	688	AAB61883	Chimeric Zcytor14
8	1888.5	50.5	575	AAB61881	Human variant Zcyt
9	1088.5	29.1	309	AAV76143	Human secreted pro
10	971.5	26.0	332	AAW25864	Human protein sequ
11	875.5	23.4	267	AAB88448	Human membrane or

12	591.5	15.8	204	AAE10920	Human gene 12 enco
13	312	8.3	617	AAV76048	Murine skin cell p
14	312	8.3	617	AAB55987	Skin cell protein,
15	310.5	8.3	667	AAU04957	Human Interleukin
16	177	4.7	864	AAW04184	Murine Interleukin
17	177	4.7	864	AAW61271	Mouse Interleukin-
18	177	4.7	864	AAW92408	Murine IL-17R prot
19	177	4.7	864	AAV99935	Murine IL-17R prot
20	177	4.7	864	AAV97130	Murine Interleukin
21	177	4.7	864	AAV97180	Murine Interleukin
22	177	4.7	864	AAV97180	Murine Interleukin
23	177	4.7	864	AAV97180	Murine IL-17R poly
24	177	4.7	864	AAV97180	Murine IL-17R poly
25	150.5	4.0	330	AAV75947	Murine skin cell s
26	150.5	4.0	330	AAV75947	Skin cell protein,
27	147.5	3.9	866	AAW55886	Human Interleukin-
28	147.5	3.9	866	AAW61272	Human Interleukin-
29	147.5	3.9	866	AAW92409	Human IL-17R prote
30	147.5	3.9	866	AAV99941	Human IL-17R prote
31	147.5	3.9	866	AAV97131	Human Interleukin-
32	147.5	3.9	866	AAV97181	Human Interleukin-
33	147.5	3.9	866	AAW03807	Human Interleukin-
34	147.5	3.9	866	AAW62066	Human IL-17R (hctL
35	147.5	3.9	866	AAV72754	Human Interleukin-
36	130	3.5	1042	ABG26706	Human Interleukin-
37	119.5	3.2	1013	AAW81163	Novel human diagno
38	112	3.0	1447	AAW68553	Mycobacterium tube
39	112	3.0	1447	AAW33498	Deleted in colorec
40	112	3.0	1447	AAW50693	Human DCC protein.
41	112	3.0	1728	AAW13144	Human UNC-40 prote
42	111.5	3.0	877	AAW28379	Deleted in Colorec
43	111	3.0	2969	AAW56442	Novel human diagno
44	108	2.9	997	ABG08088	Fragment HGJ1737 o
45	107.5	2.9	1433	ABG08373	Novel human diagno

ALIGNMENTS

RESULT 1	
AA61880	standard: Protein; 692 AA.
AA61880;	
AC	
XX	
DT	08-MAY-2001 (first entry)
XX	
DE	Human cytokine receptor Zcytor14.
XX	
KW	Cytokine receptor: Zcytor14; human; inflammation; rheumatoid arthritis;
KW	antiflammatory: gene therapy; vaccine.
XX	
OS	Homo sapiens.
XX	
PN	W0200104304-A1.
XX	
PD	18-JAN-2001.
XX	
PF	30-JUN-2000; 2000WO-US18383.
XX	
PR	07-JUL-1999; 99US-0348854.
XX	
PA	(ZYMO) ZYMOGENETICS INC.
XX	
PI	Presnell SR, Burkhead SK, Pownder SL;
XX	
DR	WPI; 2001-112618/12.
DR	N-PSDB; AAC85027.
XX	
PT	New polypeptide encoding a human cytokine receptor Zcytor14, for
PT	treating inflammation e.g. rheumatoid arthritis -
XX	
PS	Claim 2; Page 2; 112pp; English.

The invention provides a new human cytokine receptor designated zcytor14. zcytor14 can be expressed by standard recombinant methodology. The encoding nucleic acid is useful for detecting the expression of a zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be used to screen biological samples in vitro for the presence of zcytor14 proteins, polypeptides and peptides having zcytor14 activity can be administered to a subject who lacks an adequate amount of this polypeptide, for treating inflammation and conditions such as rheumatoid arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14 antibodies) can be used to treat a subject who produces an excess of zcytor14. zcytor14 nucleotide sequences can also be used to provide zcytor14 to a subject. The present sequence represents the human cytokine receptor zcytor14.

Sequence 692 AA:

Query Match 63.9%; Score 2389.5; DB 22; Length 692;
Best Local Similarity 67.0%; Pred. No. 5.4e-218;
Matches 464; Conservative 66; Mismatches 138; Indels 25; Gaps 5;

```

1 MPVSWELSLALGRNPNVSLERLMEPDTARCSLGLSCHLMDGVLCPLGSLQASAPGV 60
1 mpvswelstlalnprnvnslrlmepdtdarcsllglschlmddgvllcplgslqasapgv 60
1 mvpwflstlalnprvnslrlvpgdqdtchcpgslscrlwdsdlclpdlvppapgv 60
1 LVPFRLQTELVLRCPQKTCALFRVNVVHLAVHGHMAEPPE---AGKSDSELQESRNAS 116
1 lvpfrlqtelvlrqpoktcalfrvrvvhlavhghmaeppe---agksdselqesrnas 116
1 laptlhlqtelvlrqpoktcdclclraavhlahvhwpeedeekfgsaadsgveepnas 120
1 laptlhlqtelvlrqpoktcdclclraavhlahvhwpeedeekfgsaadsgveepnas 120
117 LQAGVYVLSFOAVPIRACALLEVOPADLYQPGOSVGSVNFECFASLQAEVQWSTTKR 176
117 lqagvyvlsfvoavpiracalleovpadlyqpgosvgsvnfecfaslqaevwsttkr 176
121 lqagvyvlsfqpvtatcalleqvvpalvqfgvgsvvycfleaalsevrlwsttqr 180
121 lqagvyvlsfqpvtatcalleqvvpalvqfgvgsvvycfleaalsevrlwsttqr 180
177 YQKELINLQOOLPDCRGLLEVRSIQSCWVLPMLNVTGDDNLTLTLDVSEEDQFSFLTYR 236
177 ykelinlqoolpdcrgllevrslqscwvlpmlnvtgddnltltldvseedqfsfltyr 236
181 ykelinhqqlp-----alpwlnvsadgdnvhlvlnveeqhffgslslynn 225
181 ykelinhqqlp-----alpwlnvsadgdnvhlvlnveeqhffgslslynn 225
237 PVPDAKSLMYKNTLNGPQNTLNTLNLVPCLCIOVWSLEPDSERVEFCFPFDPGAHRL 296
237 pvpdakslmykntlmgpqn timer ntlntlnlvpcclciovwslepdservfcpfpdp gahrl 296
226 qvqgppkprwhnltlqpgliltlnhtclvpcclqvwplpdpvrltncpfredprahqnl 285
226 qvqgppkprwhnltlqpgliltlnhtclvpcclqvwplpdpvrltncpfredprahqnl 285
297 WHIARLVLSFGWQMDAPCCLPBGKVTLCQAPDQSPCQPLVPPVPQKNATYNEQDFOL 356
297 whiarlvlsfgwqmdapccllpbgkvtlcqapdqspcqplvppvpqknatyneqdfol 356
286 wgaarllrltlqswllldpcslpaeealcwtrpaggdpcqplvpplswenlvdkylefpl 345
286 wgaarllrltlqswllldpcslpaeealcwtrpaggdpcqplvpplswenlvdkylefpl 345
357 VAGHPNLCVOVSTWKEVQLOACSMADSLGPFKDDMLLVEMKTGLNNTSVCALPESGCPPL 416
357 vaghpnlcvoovstwkevqloacsmadslgpfkddmllvemktglntsvcalpescgcppl 416
346 lkgpnlcvqvnsssklqgqclwadsipklkdvllletqpgdnsticalpescstl 405
346 lkgpnlcvqvnsssklqgqclwadsipklkdvllletqpgdnsticalpescstl 405
417 PSMASTRARLGEELLQDFRSQCMQMLMNDNGSLMACPMDKTYIHRWVVLWTLCLTLA 476
417 psmastrarlgellqdfrsqcmqlmndngslmacpmdktyihrvvvlwtlcltla 476
406 pskastreaarlgelyllqdtqsgcqlw-dddlgalwacpmkylhkrwvvlwtlcllfa 464
406 pskastreaarlgelyllqdtqsgcqlw-dddlgalwacpmkylhkrwvvlwtlcllfa 464
477 AALFFELLKTRDKRAAKRSRTRALLHSHADAGYERLVGALASLSQMPLRVAVDWLSRR 536
477 aalffellktrdkraakrsrtrallhshadagyerlvgalaslsqmplrvavdwlsrr 536
465 aalslslkldhataaagralaallysadsgferlvgalasalcqplrvavdlwstr 524
465 aalslslkldhataaagralaallysadsgferlvgalasalcqplrvavdlwstr 524
537 ELAHHGALAMFHHORRRILOEGGVVILFSPAIVAOCQOQLQTVEP---GPHALAM 593
537 elahhgalamfhhorrriloeggvvillfspaivaocqoqlqtvpe---gphalam 593
525 elsaqgqvavfhagtrqclqeggvvvllfsgavalcsewldqdvsgpghaphaftras 584
525 elsaqgqvavfhagtrqclqeggvvvllfsgavalcsewldqdvsgpghaphaftras 584
594 LSCVLPDFLQGRATRGVGVVFDGLHPDVPSPFRVAPLFLPTOLPAFLDALOGGCT 653
594 lscvlpdfqlgratrgvgvvfdglhpdvpspfrvaplflptolpafldaloggct 653
585 lscvlpdfqlgratrgvgvactdrllhpavpalftvryftlpsqpldfilgalqgprap 644
585 lscvlpdfqlgratrgvgvactdrllhpavpalftvryftlpsqpldfilgalqgprap 644
654 SAGRPADRVKXTQALRSALDSC--TSSEAPG 684
654 sagrpadvkxtqalrsaldsc--tsseapg 684
645 rsgrlgeraegvstalpaldsyfhnppgcpapg 677
645 rsgrlgeraegvstalpaldsyfhnppgcpapg 677

```

```

ID AAU29322 standard; Protein; 705 AA.
XX
AC AAU29322;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human PRO polypeptide sequence #299.
XX
KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200168848-A2.
XX
PD 20-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06520.
XX
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US03841.
PR 03-MAR-2000; 2000US-187202P.
PR 06-MAR-2000; 2000US-186968P.
PR 14-MAR-2000; 2000US-189320P.
PR 14-MAR-2000; 2000US-189328P.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000US-190828P.
PR 21-MAR-2000; 2000US-191007P.
PR 21-MAR-2000; 2000US-191048P.
PR 21-MAR-2000; 2000US-191314P.
PR 28-MAR-2000; 2000US-192655P.
PR 29-MAR-2000; 2000US-193032P.
PR 29-MAR-2000; 2000US-193053P.
PR 30-MAR-2000; 2000WO-US08439.
PR 04-APR-2000; 2000US-194449P.
PR 04-APR-2000; 2000US-194647P.
PR 11-APR-2000; 2000US-195975P.
PR 11-APR-2000; 2000US-196000P.
PR 11-APR-2000; 2000US-196187P.
PR 11-APR-2000; 2000US-196590P.
PR 11-APR-2000; 2000US-196620P.
PR 18-APR-2000; 2000US-198121P.
PR 18-APR-2000; 2000US-198585P.
PR 25-APR-2000; 2000US-199397P.
PR 25-APR-2000; 2000US-199550P.
PR 25-APR-2000; 2000US-199654P.
PR 03-MAY-2000; 2000US-201516P.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 05-JUN-2000; 2000US-209832P.
PR 28-JUL-2000; 2000WO-US20710P.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
XX
PA (GENTH ) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
XX MPI: 2001-602746/68.
XX DR N-PSDB; AAs46223.
XX
PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumours, such as prostate and breast tumours, in mammals and
PT to screen for modulators of the compounds -
XX

```


Result	3	
AA004956	standard; Protein; 705 AA.	
AA004956	24-OCT-2001 (first entry)	
Human Interleukin 17 receptor, IL-17RH2.		
Human; Interleukin-17 receptor; IL-17RH2; agonist; antagonist; PRO20040; DNA 164625-2890; systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis; diabetes mellitus; allergic disease; asthma; demyelinating disease; degenerative cartilaginous disorder; transplantation associated disease		
Homo sapiens.		
Key	Location/Qualifiers	
Peptide	1..20	
Protein	/label= Signal-peptide	
Region	21..705	
Modified-site	/label= Mature_IL_17RH2	
Region	107..112	
Modified-site	/note= "N-myristoylation site"	
Region	118..121	
Modified-site	/note= "Asn is N-glycosylated"	
Region	152..157	
Modified-site	/note= "N-myristoylation site"	
Modified-site	186..189	
Modified-site	/note= "Asn is N-glycosylated"	
Modified-site	198..201	
Modified-site	/note= "Asn is N-glycosylated"	
Modified-site	211..214	
Modified-site	/note= "Asn is N-glycosylated"	
Modified-site	238..241	
Modified-site	/note= "Asn is N-glycosylated"	
Modified-site	248..251	
Modified-site	/note= "Asn is N-glycosylated"	
Region	319..324	
Modified-site	/note= "N-myristoylation site"	
Modified-site	334..337	
Modified-site	/note= "Asn is N-glycosylated"	
Modified-site	357..360	
Modified-site	/note= "Asn is N-glycosylated"	
Modified-site	391..394	
Modified-site	/note= "Asn is N-glycosylated"	
Region	438..443	
Domain	453..473	
Region	/note= "Transmembrane domain"	
Region	516..521	
Region	/note= "N-myristoylation site"	
Region	552..555	
Region	/note= "cAMP/GMP-dependent protein kinase phosphorylation site"	
Region	583..586	
Region	/note= "Glycosaminoglycan attachment site"	
Region	612..617	
Region	/note= "N-myristoylation site"	
Region	692..697	
Region	/note= "N-myristoylation site"	
Region	696..701	
Region	/note= "N-myristoylation site"	
Region	700..705	
Region	/note= "N-myristoylation site"	
W0200146420-A2.		
28-JUN-2001.		

20-DEC-2000; 2000MO-US34956.
 23-DEC-1999; 99US-0172096.
 30-DEC-1999; 99MO-US31274.
 11-JAN-2000; 2000US-0175461.
 18-FEB-2000; 2000MO-US04341.
 02-MAR-2000; 2000MO-US05841.
 21-MAR-2000; 2000US-0191007.
 02-JUN-2000; 2000MO-US07542.
 22-JUN-2000; 2000US-0213087.
 22-AUG-2000; 2000US-0644848.
 24-AUG-2000; 2000MO-US23828.
 24-OCT-2000; 2000US-0242837.
 10-NOV-2000; 2000MO-US30873.
 28-NOV-2000; 2000US-0253646.
 01-DEC-2000; 2000MO-US32678.

(GETH) GENENTECH INC.

Chen J, Filvaroff E, Fong S, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Li H, Hillan KJ, Tamas D, Van Lookeren M, Vandlen KL, Watanabe CK, Williams PM, Wood WI, Yansura DG;
 WPI; 2001-451708/48.
 N-PSDB; AAS09515.

Novel PRO polypeptides homologous to interleukin-17, useful for the diagnosis and treatment of immune related disease e.g. rheumatoid arthritis and diabetes
 Claim 10; Fig 14; 188pp; English.

The sequence is PRO20040 which is the human interleukin 17 receptor, IL-17R2, encoded by DNA 164625-2890. A composition containing ant/agonists to the PRO polypeptides or individual components are useful for treating a mammal with an immune related disease, e.g. systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an idiopathic inflammatory myopathy, Sjogren's syndrome, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal disease, a demyelinating disease, an autoimmune or immune-mediated skin disease, contact dermatitis, an allergic disease e.g. food hypersensitivity, asthma, a transplantation associated disease, or a chronic inflammatory demyelinating polyneuropathy. Treating a degenerative cartilaginous disorder comprises administering a PRO1031 or PRO1122 polypeptide agonist, or antagonist to the mammal. Numerous examples of the diseases and disorders are given in the specification.

Sequence 705 AA:

Query Match 63.5%; Score 2376; DB 22; Length 705;

Best Local Similarity 65.9%; Pred. No. 1,1e-216; Matches 466; Conservative 66; Mismatches 135; Indels 40; Gaps 7;

1 MVSFSLSLALGRNVVSLERLMEPDTRARCSLGLSCHLMDGVLCIPGSLQSNAPGV 60
 1 mpvfwllslalgrnpvslervlpqdathepsglscrlwdsdlcllpgtlvpaqpv 60
 61 LVTPLQRTLVLRCPQKTCALNRVVVHLAVHHAPEE-----AGSDSELOESRRNAS 116
 61 laptlqlrtlvlrqetdclclvaavhavlghweepdeekfsgaadvgeepnas 120
 117 LQAQVLSQAPVPIACALLEVOPADLVOPGSGVSAVFECFASLGAAYVIMSYTKR 176
 121 lqagvlsfqapvpiatrcavllvqvpaalvqfsgvsgvavycfeaaqsevrliwstqpr 180
 177 YQKELNLTQOALPDCRGLVVDSTQSCWVLPFWLNVSTGDNVLTLLDVSEEDPFSFLYLR 236
 181 yekelnltqqlp-----alpwlnvsagdnvhlvlnveeqhfglslywn 225

QY 237 PVPDALKSLMKNTLTGQONTLNTDVLVPCLCIOVWSLEPDSERYEFCFPEDPGAHRNL 296
 Db 226 qvqgppkprvwhkntltpqiltlnhtldvpcldciqvwplepdsrvtnicpfredprahqnl 285
 QY 297 WHIARLRVLSFGVQVQLAPCCLPKVTLCQWAPDQSPCOPLVPVPOKRNATVNEPDQFL 356
 Db 286 wgaarlrlltlqsvllapcslpeaaalcwarpagdpqplvplswenvlvdkvlefp 345
 QY 357 VAGHPNLGVQVSTWKEVQVQLACSWADSLGPFKDDMLVEMKGTNLNTSCALPESGCTPL 416
 Db 346 lkgpnlcvgvnsekqlqgclwadsigpklkdvlllletrgpqnrlscalpepsctsl 405
 QY 417 PSMSTRRAALGEEILLDDFHSQCQWLNNDNMGSLACPMKDYHRRKVLVWLACLLA 476
 Db 406 pskasttraarlgeyllqdlqdsqgclqlw-dddlgalwacpmkylhkrvalvwlacllfa 464
 QY 477 AALFFELLKKRRK-----AARGSTRALLNSADGAGYERLVGALASALS 522
 Db 465 aalsllllllkkdhakgwylrlkqdvrgsaarg-raalllvsaddsgterlvgalasalc 523
 QY 523 QMFLRAVADLMSRRRLSAHGALAMFHHRRLLOEGVVLTFSPAAVAQCCOMIQLQTV 582
 Db 524 qplrvavdlwstrrelsaqgpvaawfhagrtqlqegvrvllfsgvaalacsewldqgvs 583
 QY 583 EP---GPHQALAMLSGVLPDFLOGRATGRVGVYFDGLHPDVSPPRVAPLSLPMQ 639
 Db 584 gpgahbphdaftraslscvlpdflqgrapsyvgacfdllhpdaavalfrtvvflfpq 643
 QY 640 LPAFLPALOGCGSTSGAGRPADRVETRVQALRSALDSC--TSSSEAG 684
 Db 644 lpdflgalqgraprsgrlqgraeqvaralqpaldsyfhppqtpapq 690

RESULT 4

AAB87606
 ID AAB87606 standard; Protein: 705 AA.

XX AAB87606;

DT 15-MAY-2001 (first entry)

DE Human PRO20040.

KW Human; PRO protein; mapping.

OS Homo sapiens.

PN WO200116318-A2.

XX 08-MAR-2001.

PD 24-AUG-2000; 2000MO-US23328.

PF 01-SEP-1999; 99MO-US20111.

PR 15-SEP-1999; 99MO-US21090.

PR 07-DEC-1999; 99US-0169495.

PR 09-DEC-1999; 99US-0170262.

PR 11-JAN-2000; 2000US-0175481.

PR 18-FEB-2000; 2000MO-US04341.

PR 18-FEB-2000; 2000MO-US04342.

PR 22-FEB-2000; 2000MO-US04414.

PR 01-MAR-2000; 2000MO-US05601.

PR 03-MAR-2000; 2000US-0187202.

PR 25-APR-2000; 2000US-0199397.

PR 22-MAY-2000; 2000MO-US14042.

PR 05-JUN-2000; 2000US-0209832.

(GETH) GENENTECH INC.

Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;

WPI; 2001-183260/18.

QY 237 PVPPDALKSLWYKNLTGPQNTLNTHTDLYPCLCIQWWSLEPDSERVEEFCPRDPGAHRNL 296

23/ PVPDAKSLMYRNLTPQNTLNLHTDVPCLCIQWWSLEPDSERVEFCPPRRDPGAHRTL 290

226 qvqgppkprvkhnltpqitlnthtlvpclciqvwplpdpstvrtncipfredprahnl 285
 297 WHIARLRLVSPGWOLADAPCCIPGKVTLCQWADPQSPCLVPEVPQKNATVNEPQDFOL 356
 286 wgaarlrltlqswlldapcsipaeealcwarpagpcqplvpplswenvtvdvlefpj 345
 357 VAGHPNLCVOYSTWEKVOLOACSNADSLGPFKDMILVEMKGTGLNNTSVCALEPSCGTP 416
 346 lkgpnclcvgnsskqlqecclwadsipklkdvlllettrpqdnrlcalpescstl 405
 417 PSMASTRARAGEELIDPFRSHQCQGLANDDNMGSIMACPMKYIHRRWLVWLACILLA 476
 406 pskastraarlgeylldqdgscqlw-dddglawacpmkylhkrwlvwlacllla 464
 477 ALFFELLKKDRK-----AAROSRTALLHSADGACGERLVGALASALS 522
 465 aalslillllkkdhakglwrltkqdvtrsgaarg-raaillysadsgferlvgalasalc 523
 523 QMPLRVAVDLSRRRLSHAGLAWFHRRRLIQEGGVILLFSPAAYAOCCQMLQIQTV 582
 524 qlplrvavdlvstrrelsqgprvawfhagtrqlqeggvvllfsgavalcsewldqgvs 583
 583 EP---GPHDALAAMLCVLPDFOGRATGRVYGVFDGLHPDSVPSFRVAPLFSLPTQ 639
 584 gpgahghdetralscylpdtflggrapgsyvgacfdrlhhpavpalftvplvftlpsq 643
 640 LPAFLDALOGGCGSTSGRPAADREVERTOALRSALDSC--TSSSEAPG 684
 644 lpfdfgalqgpraprgrlqeraeqvsralqpaldeyfhnpptpapg 690
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

CC arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14
 CC antibodies) can be used to treat a subject who produces an excess of
 CC zcytor14. zcytor14 nucleotide sequences can also be used to provide
 CC zcytor14 to a subject. The present sequence represents a chimeric
 CC zcytor14 protein.
 XX
 XX
 Sequence 675 AA;
 Query Match 61.9%; Score 2314; DB 22; Length 675;
 Best Local Similarity 65.5%; Pred. No. 7.9e-211;
 Matches 454; Conservative 63; Mismatches 134; Indels 42; Gaps 6;
 1 MPVSMFLSLAGRPVYVSLERLMEPOFATCSGLSCHLMGDVLCGLGSLQASAGPV 60
 1 mpywflslalgrpsvslsterlvpgqdaclncspjlsclwdsdlclpgdlvpaqpv 60
 61 LVPRLOTELVRCPQKTDICALRVVWVLAHGHMAEPD---AGKSDSELOESRNAS 116
 61 laphlqtelvllrqketcdclrvavnlavhghwepeedeekfggaadsyveepnas 120
 117 LQAVVLSFOAIPYAKCALLEVQPADLVQPGQSVGSAVDECEASLGAETQWSTYKPR 176
 121 lqavvlstfgayptarcvllvqvpaalvqfgsvgsyvvydcfeaalgservlswytlqpr 180
 177 YQKELNTQOOLPDORGLEVRDSIQSCWVLPWLVNSTRDGNVLLTLVSEEDPSFLYLX 236
 181 yekelntqqlp-----alpwlvnsadgdnvhlvlvnsqehgslslyvn 225
 237 PVDPALKSLWYKNTLGPONITLNTHTLVPCICIQVWSTLEPDSERVECEPREDPAGHRL 296
 226 qvqgppkprvkhnltpqitlnthtlvpclciqvwplpdpstvrtncipfredprahnl 285
 297 WHIARLRLVSPGWOLADAPCCIPGKVTLCQWADPQSPCLVPEVPQKNATVNEPQDFOL 356
 286 wgaarlrltlqswlldapcsipaeealcwarpagpcqplvpplswenvtvdvlefpj 345
 357 VAGHPNLCVOYSTWEKVOLOACSNADSLGPFKDMILVEMKGTGLNNTSVCALEPSCGTP 416
 346 lkgpnclcvgnsskqlqecclwadsipklkdvlllettrpqdnrlcalpescstl 405
 417 PSMASTRARAGEELIDPFRSHQCQGLANDDNMGSIMACPMKYIHRRWLVWLACILLA 476
 406 pskastraarlgeylldqdgscqlw-dddglawacpmkylhkrwlvwlacllla 464
 477 ALFFELLKKDRKRAAGSFTALLHSADGACGERLVGALASALSQMLRAVADLMSRR 536
 448 aalslillllkkdhakaargraaillysadsgferlvgalasalcqlplrvavdlwstr 507
 537 ELSAHGALAMFHORRRILIOGGGVILLFSPAAYAOCCQMLQIQTV 582
 508 elsaqgprvawfhagtrqlqeggvvllfsgavalcsewldqgvsqgprahghdetrals 567
 594 LSCVLPDFOGRATGRVYGVFDGLHPDSVPSFRVAPLFSLPTQOLPFLDALOGGCGST 653
 568 lscvlpdfllggrapgsyvgacfdrlhhpavpalftvplvftlpsqldfllgalsqprap 627
 654 SAGRPADREVERTOALRSALDSC--TSSSEAPG 684
 628 rsgrlqeraeqvsralqpaldeyfhnpptpapg 660
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

Claim 2; Page 107-109; 112pp; English.

New polypeptide encoding a human cytokine receptor zcytor14, for treating inflammation e.g. rheumatoid arthritis -

WPI; 2001-112618/12.

Presnell SR, Burkhead SK, Powderer SL;

(ZYMO) ZYMOGENETICS INC.

07-JUL-1999; 99US-0348854.

30-JUN-2000; 2000MO-US18383.

18-JAN-2001.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

WO200104304-A1.

KW antiinflammatory; gene therapy; vaccine.
 XX
 OS Homo sapiens.
 PN WO200104304-A1.
 XX 18-JAN-2001.
 PD
 PF 30-JUN-2000; 2000MO-US18383.
 PR 07-JUL-1999; 9905-0348854.
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 PI Presnell SR, Burkhead SK, Powder SL;
 DR WPI: 2001-112618/12.
 XX
 PT New polypeptide encoding a human cytokine receptor zcytor14, for
 PR treating inflammation e.g. rheumatoid arthritis -
 XX
 PS Claim 2: Page 102-104; 112pp: English.
 XX
 CC The invention provides a new human cytokine receptor designated zcytor14.
 CC zcytor14 can be expressed by standard recombinant methodology. The
 CC encoding nucleic acid is useful for detecting the expression of a
 CC zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be
 CC used to screen biological samples in vitro for the presence of zcytor14.
 CC proteins, polypeptides and peptides having zcytor14 activity can be
 CC administered to a subject who lacks an adequate amount of this
 CC polypeptide, for treating inflammation and conditions such as rheumatoid
 CC arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14
 CC antibodies) can be used to treat a subject who produces an excess of
 CC zcytor14. zcytor14 nucleotide sequences can also be used to provide
 CC zcytor14 to a subject. The present sequence represents a chimeric
 CC zcytor14 protein.
 CC
 SQ Sequence 688 AA:
 Query Match 61.5%; Score 2300.5; DB 22; Length 688;
 Best Local Similarity 64.5%; Pred. No. 1.6e-209;
 Matches 456; Conservative 63; Mismatches 131; Indels 57; Gaps 8;
 QY 1 MPVSWFLSLALGRNPVYVLERLMEPODTRARCSIGLSCHMDGDLCLPGSLQASGPV 60
 DB 1 MPVWFLSLALGRNPVYVLERLMEPODTRARCSIGLSCHMDGDLCLPGSLQASGPV 60
 QY 61 IVPTRLOTETELRCPKOTDCALRYRVVYVHLAVHGMAREE-----AGKDSLEQESRNAS 116
 DB 61 IAPTHLGTETELRCPKOTDCALRYRVVYVHLAVHGMAREE-----AGKDSLEQESRNAS 116
 QY 117 LQAQVVLSEFQAYPIARCALLEVOPADLVOPGSGSAVFDFEASLGAEOVMSTYKPR 176
 DB 121 LQAGVVLSEFQAYPIARCALLEVOPADLVOPGSGSAVFDFEASLGAEOVMSTYKPR 176
 QY 177 YQKELNLTQQLPDCRGLERDSTQSCVNLVWLVNSTDGNTLLTLDVSEODFSFLLYLR 236
 DB 181 YQKELNLTQQLPDCRGLERDSTQSCVNLVWLVNSTDGNTLLTLDVSEODFSFLLYLR 236
 QY 237 PVPDALKSLMYKNLGTGPONTLNTHTDLPVPCICIQWMSLEPSEEREFCEFPREDPGAHRL 296
 DB 226 GYQGPDKPRWKNLTGPQILHTDLPVPCICIQWMSLEPSEEREFCEFPREDPGAHRL 296
 QY 297 WHIARLVLSPGVWQADAPCLPGKVTLCMQAPDOSPCOPLVPVPVOKNATVNEPQDFOL 356
 DB 286 WGAARLVLSPGVWQADAPCLPGKVTLCMQAPDOSPCOPLVPVPVOKNATVNEPQDFOL 356
 QY 357 VAGHPMLCVQVSTWVEVOAQCSWADSLGEPKDMILVEMKTLGNLNTSVCALEPSCPTL 416
 DB 339 VAGHPMLCVQVSTWVEVOAQCSWADSLGEPKDMILVEMKTLGNLNTSVCALEPSCPTL 416
 QY 417 PSMASRAARLAGEELLQDFRSHQCMQMLNDNDMGSLVACPDKXTHRRVYLVYLACLILA 476

DB 389 PSKASTRAARIGEYLLGLDLSGQCQLW-dddLGLAWACPMKYIHKRWALVYLACILFA 447
 QY 477 AALFFFLTKKDRRK-----AAGSRFALLHSADGAGYERLVCALASALS 522
 DB 448 AALSLLILKKDKHAKYRLLKQDYSQAARG-raallllysdsgferllygalaalc 506
 QY 523 QMPLRAVADLMSRRRLSAGALAMFHHQRRILOEGVYLLFSFAVACQOQMLQOTV 582
 DB 507 QPLRVAVDLSRRRLSAGALAMFHHQRRILOEGVYLLFSFAVACQOQMLQOTV 582
 QY 583 EP---GPHDALAMLSCYLPDFLOGRANGRYGVTFDGLHPDSVSPFRVADLSLPTQ 639
 DB 567 GPGAHGPHDAFASLSCVLPDFLOGRANGRYGVTFDGLHPDSVSPFRVADLSLPTQ 639
 QY 640 LPAFLDAGCGSTRSAGRPADRVERTQALRSALDSC--TSSEARG 684
 DB 627 LPDFFLGLQGPRAPRSGRIGERAEQVSRALGPALDSFYHPGLTPAPG 673
 RESULT 8
 ID AAB61881 standard; Protein: 575 AA.
 AC AAB61881;
 DT 08-MAY-2001 (first entry)
 DE Human variant zcytor14 protein zcytor14-1.
 XX
 KW Cytokine receptor; zcytor14; human; inflammation; rheumatoid arthritis;
 KW antiinflammatory; gene therapy; vaccine; variant; zcytor14-1.
 OS Homo sapiens.
 XX
 PN WO200104304-A1.
 XX 18-JAN-2001.
 PD
 PF 30-JUN-2000; 2000MO-US18383.
 PR 07-JUL-1999; 9905-0348854.
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 PI Presnell SR, Burkhead SK, Powder SL;
 DR WPI: 2001-112618/12.
 DR N-PSDB; AAC85029.
 PT New polypeptide encoding a human cytokine receptor zcytor14, for
 PR treating inflammation e.g. rheumatoid arthritis -
 XX
 PS Disclosure: Page 2-3; 112pp: English.
 XX
 CC The invention provides a new human cytokine receptor designated zcytor14.
 CC zcytor14 can be expressed by standard recombinant methodology. The
 CC encoding nucleic acid is useful for detecting the expression of a
 CC zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be
 CC used to screen biological samples in vitro for the presence of zcytor14.
 CC proteins, polypeptides and peptides having zcytor14 activity can be
 CC administered to a subject who lacks an adequate amount of this
 CC polypeptide, for treating inflammation and conditions such as rheumatoid
 CC arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14
 CC antibodies) can be used to treat a subject who produces an excess of
 CC zcytor14. zcytor14 nucleotide sequences can also be used to provide
 CC zcytor14 to a subject. The present sequence represents a variant of
 CC the human cytokine receptor zcytor14, designated zcytor14-1. This
 CC variant is a truncated form of the receptor polypeptide and lacks
 CC amino acid residues 1-113 of zcytor14.
 CC
 SQ Sequence 575 AA:

Query Match 50.5%; Score 1888.5; DB 22; Length 575;
Best Local Similarity 63.8%; Pred. No. 1.9e-170;
Matches 379; Conservative 52; Mismatches 110; Indels 53; Gaps 7;

110 QESNASTQANVLSFQVPIKALLEVQVADLVQPGQVSGASVDFCEBASLGAEOI 169
1 eepnaaiaqavvlsifqyplacvllvqvpaaivqfsgsvsvvycfcaalgsevr 60

170 MSTRPKRYQKELNTQOQPCDGLVQSDIOSQVLPMLNVSDGQVILTLQVSEDF 229
61 wstypqpyekelntqgqg-----alpwlnvsadgdnvhlvlnvseeght 105

230 SFLLYLRPPDALSLWYKNTLTPONTLNHTDVLVPCILQVWSLEPDSERVECPFERD 289
106 gslvynvqvgpprpvkhnltpqlltlnhclvpcclqvwplepdsrvrtlcpfred 165

290 PGARNIMHTARLVLSGCVQQLDAPCCLPKVTLCWQAPDQSPQPLVPVPQKNATVN 349
166 ptnaqalwqaarlltlltqsvllldapcslpaaaalcwrapgqpcqplvpplswenlvtd 225

350 EPDQFQVAGHPNLCVQSTWEKVOLOACSNADSLGPFKDMLLVEKKTGLNNTSCALE 409
226 -----vnsekqlgeclwadsipkladvllletrgpqdnrlcale 268

410 PSCCTLPSPMASTRARLGEELLQDFRSHQCNQMLNDNMKSLMACPMKYIHRRWLVW 469
269 psgcslspskasttraarlgelyllqdlsgqclqvw-dddgalwacpmkyihkrwalvw 327

470 LACLLAAALFEFLTKDRRK-----AARGSRALLHSDAGATERLVG 515
328 lacllfaaaalstlllkdkhkgwlrllkqdvrgsaaarg-raalllyssddsgferlvq 386

516 ALASALSQMPLRVAVDMSRRELNSHAGALWFHQRRLIQEGVVLTSPPAVAACQ 575
387 alasalqlrlvavdlvsrrrelsqgvrwfhagrrlqlqegsvvlllspgvalcse 446

576 WQLQVTEP---GRPDALAAULSCVLPFLOGRATRGVGVVFDGLHPSVSPFRVAP 632
447 wldqvgvsgpabghdatriascvlpdlfgrtapsyvgacfdrlhlpdaavpalfrtvp 506

633 LFSLPRLPFLDALQGGCSTAGRAPRADERVETQALRSALDSC--TSSEAPG 684
507 vflfslpdlpdlfslgqlqpraprgslrgetaegvralpaldsyflppgtpapg 560

RESULT 9
AY76143
AAV76143 standard; Protein: 309 AA.
AAV76143:
23-MAR-2000 (first entry)
Human secreted protein encoded by gene 20.

Human: secreted protein; cancer; tumour; developmental abnormality;
foetal deficiency; blood disorder; immune system disorder; inflammation;
autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
therapy; chromosome 3.

Homo sapiens.
MO9958660-A1.
18-NOV-1999.
06-MAY-1999; 99MO-US09847.
12-MAY-1998; 98US-0085093.

PR 12-MAY-1998; 98US-0085094.
PR 12-MAY-1998; 98US-0085105.
PR 12-MAY-1998; 98US-0085180.
PR 18-MAY-1998; 98US-0085906.
PR 18-MAY-1998; 98US-0085920.
PR 18-MAY-1998; 98US-0085921.
PR 18-MAY-1998; 98US-0085922.
PR 18-MAY-1998; 98US-0085923.
PR 18-MAY-1998; 98US-0085924.
PR 18-MAY-1998; 98US-0085928.
PR 18-MAY-1998; 98US-0085925.
PR 18-MAY-1998; 98US-0085927.

(HUMA-) HUMAN GENOME SCI INC.
Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;
Lafleur DW, Endress GA, Ebner R.
WPI: 2000-06296/05.
N-PSDB: AA65269.

New isolated human genes and the secreted polypeptides they encode,
useful for diagnosis and treatment of e.g. cancers, neurological
disorders, immune diseases, inflammation or blood disorders

Claim 11: Page 372-373; 475pp; English.

AA65250 to AA65350 represent 97 isolated human secreted protein genes.
AAV76124 to AAV76223 are the secreted proteins encoded by the 97 human
genes. The gene encoding this protein was found to be on chromosome 3.
The genes and their corresponding secreted polypeptides are
useful for preventing, treating or ameliorating medical conditions,
e.g. by protein or gene therapy. Also pathological conditions can be
diagnosed by determining the amount of the new polypeptides in a sample
or by determining the presence of mutations in the new genes. Specific
uses are described for each of the 97 genes, based on which tissues they
are most highly expressed in, and include developing products for the
diagnosis or treatment of cancer, tumours, developmental abnormalities,
and foetal deficiencies, blood disorders, diseases of the immune system,
autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
disorders, digestive/endocrine disorders, infections and AIDS. The
polypeptides are also useful for identifying their binding partners.
The sequences shown in AAV76224 to AAV76424 represent fragments of the
secreted proteins.

Sequence 309 AA:
SQ

Query Match 29.1%; Score 1088.5; DB 21; Length 309;
Best Local Similarity 71.0%; Pred. No. 9.3e-95;
Matches 211; Conservative 21; Mismatches 60; Indels 5; Gaps 2;

1 MPVSWFLSLALGNPPVVSLERLMEPDTRACSLGSLMDGVDVLCPLGSLQSAAPV 60
1 mprvswflslalgsprvslsterlvpgpdaticspglscrldsdllcpgdlvpppgv 60

61 LVPRFLQTELVLCRQKDTQCLARVRVVHVLAVHGMABEPE---AGKSDSELQESNMAS 116
61 lvprflqtevlvrcqkdtclclrvxvhlavhgmwepeedekfgaadlgyceprnas 120

117 LQNOVLSFQVPIKALLEVQVADLVQPGQVSGASVDFCEBASLGAEOI 176
121 lqnvvlsfqaqparcvllvqvpaaivqfsgsvsvvycfcaalgsevrilwstypqr 180

177 YQKELNTQOQPCDGLVQSDIOSQVLPMLNVSDGQVILTLQVSEDFCEBASLGAEOI 236
181 ykexhntqgldpcrglevmvsipscevalpwlnvsadgdnvhlvlnvseeghtgslvyn 240

237 PVPDALSLWYKNTLTPONTLNHTDVLVPCILQVWSLEPDSERVECPFERDPGAH 293
1 pvpdalslwnltpgontlnhtdlvpcilqvwslpdservfcpfrrdpgha 293

Db 241 qvqgpkprwhknltpqitltlnhtclvpcldqvwpdpdsvrtpsapgrtp-ah 296

RESULT 10

AAAM25864 standard; Protein: 332 AA.

AAAM25864:

16-OCT-2001 (first entry)

Human protein sequence SEQ ID:1379.

Human; cancer; ulcer; HIV infection; human immunodeficiency virus; antihistaminic; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac; central nervous system; vitreous; anti-HIV; fungicide; antitumor; cardiovascular; antianemic; anaemia; antiaggregant; haemostatic; vulnery; antilucer; osteopathic; eczema; dermatological; antiallergic; antiasomatic; antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; infection; immunostimulant; gene therapy; antisense therapy; vaccine; inflammation; antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis; cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity; genetic disease; haematopoietic disorder; platelet disorder; asthma; thrombocytopaenia; osteoporosis; severe combined immunodeficiency; allergic rhinitis; diabetes; multiple sclerosis; depression; Alzheimer's disease; Parkinson's disease; neurodegenerative disorder; neurological disorder.

Homo sapiens.

WO200153455-A2.

26-JUL-2001.

22-DEC-2000; 2000WO-US35017.

23-DEC-1999; 99US-0471275.

21-JAN-2000; 2000US-0488725.

25-APR-2000; 2000US-0552317.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI: 2001-457603/49.

N-PSDB: AAH99805.

Claim 20; Page 283; 1217PP; English.

AAH99166 to AAH99904 encode the human proteins given in AAAM25225 to AAAM25963. The proteins can have activities based on the tissues and cells they are expressed in, such as: antitumorigenic; antirheumatic; antiallergic; immunosuppressive; antibacterial; endocrine; cardiac; central nervous system; vitreous; anti-HIV; fungicide; antitumor; cardiovascular; antianemic; antiaggregant; haemostatic; vulnery; antilucer; osteopathic; dermatological; antiallergic; antiasomatic; antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and

CC neurological disorders.

XX Sequence 332 AA;

Query Match 26.08; Score 971.5; DB 22; Length 332;
Best Local Similarity 69.18; Pred. No. 1.4e-83;
Matches 188; Conservative 23; Mismatches 56; Indels 5; Gaps 2;

QY 1 MPVSWFLSLALGRNPVSVLERLMEPODTARCSIGLSCHLMQDVLCIPGSLQSAAGPV 60
DB 2 MPVWFLSLALGRNPVSVLERLMEPODTARCSIGLSCHLMQDVLCIPGSLQSAAGPV 61
QY 61 LVPTRIOTELVLRCPQKTDCAIRRVVHLAVGHMAEPEE---AGKSDSELEQESRNAS 116
DB 62 LAPTHTQTELVLRCPQKTDCAIRRVVHLAVGHMAEPEE---AGKSDSELEQESRNAS 121
QY 117 LQAGVYLSFQAPYFARCALLLEVOPADLVOPQSGSANVFQCFEASLGAELVQIWTSTKPR 176
DB 122 LQAGVYLSFQAPYFARCALLLEVOPADLVOPQSGSANVFQCFEASLGAELVQIWTSTKPR 181
QY 177 YQELNLTQQLPDCGLEVRDSIOSCVLPMVNSTDGNVLTLPVSEEDPSFLLYLR 236
DB 182 YQELNLTQQLPDCGLEVRDSIOSCVLPMVNSTDGNVLTLPVSEEDPSFLLYLR 241
QY 237 PVPDALKSLWYKNLTGPQNTLTNHTDLPVPCIC 268
DB 242 qvqgpkprwhknltpqitltlnhtclvpcldqvwpdpdsvrtpsapgrtp-ah 296

RESULT 11

AAAB8448 standard; Protein: 267 AA.

AAAB8448:

23-MAY-2001 (first entry)

Human membrane or secretory protein clone PSEC0233.

Human; secretory protein; membrane protein; vaccine; gene therapy;

rheumatoid arthritis; diabetes.

Homo sapiens.

EP1067182-A2.

10-JAN-2001.

07-JUL-2000; 2000EP-0114090.

08-JUL-1999; 99EP-0194179.

11-JAN-2000; 2000JP-0118775.

02-MAY-2000; 2000JP-0183766.

(HELI-) HELIX RES INST.

Ota T, Isegaki T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;

WPI: 2001-093869/11.

N-PSDB: AAF93875.

Nucleic acids encoding secretory proteins/membrane proteins, useful in gene therapy or as candidate target molecules in drug development -

Claim 1; SEQ ID 264; 609pp + CD ROM; English.

This invention relates to nucleic acid sequences AAF93744 - AAF93916 which encode human secretory or membrane proteins represented by

CC AAAB8448 - AAAB8449. Included in the invention are primers

CC AAF93917 - AAF94295 and AAF92232 - AAF92235 which are used to isolate the

CC cDNA sequences of the invention. The invention also includes methods for

CC the production of antibodies directed against the proteins, and cDNA

XX AAY76048;
 AC 27-MAR-2000 (first entry)
 DT XX
 DE Murine skin cell protein, SEQ ID NO:303.
 XX
 KM Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
 KM embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
 KM secreted; transmembrane; inflammation; cancer; neurological disease;
 KM angiogenesis; tumour vascularisation; growth disorder;
 KM developmental disorder; skin wound; hair follicle disorder;
 KM anti-inflammatory; cytostatic; neuroprotective; vulnery.
 XX
 OS Mus sp.
 XX
 PN W09955865.A1.
 PD 04-NOV-1999.
 XX
 PF 29-APR-1999; 99WO-N200051.
 XX
 PR 29-APR-1998; 98US-0069726.
 PR 09-NOV-1998; 98US-0188930.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 RI Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
 XX
 DR WPI; 2000-072177/06.
 DR N-PSDB; AAZ61753.
 XX
 PT Novel polynucleotides useful for the treatment of various conditions
 XX including wounds and cancer -
 PS
 PS Claim 4; Page 179-180; 235pp; English.
 XX
 CC The invention relates to novel nucleic acid sequences derived from rat
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
 CC cells. Polypeptides of the invention may be used to treat inflammation,
 CC cancer and neurological diseases. The proteins may be used to stimulate
 CC the growth and motility of keratinocytes, to inhibit the growth of
 CC cancer cells, to modulate angiogenesis and tumour vascularisation, to
 CC modulate skin inflammation, to modulate epithelial cell growth and to
 CC inhibit binding of HIV-1 to leukocytes. The invention may also be used
 CC to treat growth and developmental defects, skin wounds and hair follicle
 CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded
 CC by cDNA sequences derived from several mouse, rat or human skin cell
 CC types. Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and
 CC AAY76119 are proteins with an N-terminal signal sequence, indicating
 CC that they are secreted. Sequences AAY75986-Y75989, AAY76061-Y76071,
 CC AAY76106-Y76109 and AAY76121-Y76122 are proteins with one or more
 CC putative transmembrane domains.
 CC
 SO Sequence 617 AA;

Query Match 8.3%; Score 312; DB 21; Length 617;
 Best Local Similarity 23.9%; Pred. NO. 2, 1e-20;
 Matches 134; Conservative 66; Mismatches 224; Indels 136; Gaps 21;

QY 175 PRYQKELNLTQOOLPDKRLV-----RDSIOSC--WVLPMLNVSFDGDNVLLTLVDVSEE 226
 DB 137 pelsfdl-----lpevqavrvllpaggkasyrlcyqwaletedlsspfdt----- 181
 QY 227 QDSFLLYLPVPAKLSIMYKNLTGPNITLNMHDVPCICIGVWNSLEPDSERVERECRF 286
 DB 182 -----qkivsghtvdlpyefllipcmcleasyldgedvyrkkcpf 221
 QY 287 RENDPGAH-RNIMHARLNVLSPG---WVOLDAPCCLPGRVTLQWAPDQPCQPLVPPVP 342
 DB 222 qswpeaygsdfwgsifrfcdysqhnqvmaltlrcplkleaslcwrgdpltpcctl----- 276

QY 343 OKNATVNEPQDFQVAG---HPNLCVQVSTWEKVQLQAC-----SMADSLGPFKDDM 391
 DB 277 -pnatagesegyllenvdllhpqlcfkfs--fenshvecpbgsgslpswtvsmtd-qaqg 333
 QY 392 LLYEMKRTGINTNSVCLAEPSGC---TPLEPMASTRAARLGEELLQD-----FRSHQCMQL 443
 DB 334 ltlhfsrtlyafsaasdpjlgpdtmmpvysisqgsvpvtldllpflrgencllv 393
 QY 444 WNDWMGSLMACPMOKYIHRRWLV-----WACLILAAALFFLLKKDRRAAAGS 496
 DB 394 wrsd-----vlfawhvlcpddapylptglll-----tslsgqrt 427
 QY 497 RTALLHSADGAGYERLVGALASALSOM---PLRAVADLMSRRELSAHGALAMFHOHR 553
 DB 428 rpylllhaadseaqrllvgaleellrtalggydvrvldwvgtvharipjlpwlaarer 487
 QY 554 ILOEGGVILFSPAIVAOCCQMLQVTEPPGPHDALAMLSCVLPFLQGRATGRYGV 613
 DB 488 varegvtvllwncagps-----tacsqdpqaaslrlclcaaprl-----lla 531
 QY 614 YFDGLHPDSVSPFRVAVLFLPQLPAPFLDALQGCSTSA-----GRPADRYERTQ 667
 DB 532 yftrlcakgdlprlralprylrrdlprllraldaqpataaswshlgakrciknrleq 591
 QY 668 A-----LRGALDSCSTSSSEAP 683
 DB 592 chlleaakddygstnsp 611

RESULT 14

AAB55987 standard; Protein; 617 AA.

AAB55987;

08-MAR-2001 (first entry)

Skin cell protein, SEQ ID NO: 303.

Mouse; skin cell; cytostatic; antiinflammatory; anti-HIV;

keratinocyte growth stimulation; vulnery; immunomodulatory; vaccine;

inflammation; neurological disease.

Mus sp.

W020006984-A2.

23-NOV-2000.

15-MAY-2000; 2000WO-N200075.

14-MAY-1999; 99US-0312283.

(GENE-) GENESIS RES & DEV CORP LTD.

Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD, Murison JG;

WPI; 2001-007495/01.

N-PSDB; AAC99686.

New isolated polynucleotide used in the identification of genetic

disorders and encoding polypeptides used for treating inflammatory

disease, cancer and neurological diseases -

Claim 4; Page 244-245; 352pp; English.

The present sequence is a polypeptide which is expressed in

mammalian skin cells. The polypeptide is useful for stimulating

keratinocyte growth and motility, inhibiting the growth of cancer cells,

modulating angiogenesis, inhibiting angiogenesis and vascularisation of

tumours, modulating skin inflammation, stimulating the growth of

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2002, 19:42:54 ; Search time 72.7 Seconds

(without alignments)
1660.942 Million cell updates/sec

Title: US-09-899-471-5
Perfect score: 3741
Sequence: 1 MPVSWFLSLALGRNPVVS.....SSAPGCCCEMDLGPCTTLE 698

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2629	70.3	567	11	099J43 mus musculu
2	1786.5	47.8	538	4	09BR97 O9br97 homo sapien
3	177	4.7	864	11	060943 mus musculu
4	147.5	3.9	866	4	043844 homo sapien
5	147.5	3.9	866	4	096R46 homo sapien
6	119.5	3.2	1013	16	053499 mycobacteri
7	114	3.0	3021	12	068870 hepatitis c
8	112	3.0	1445	11	063155 rattus norv
9	111.5	3.0	2910	11	055225 mus musculu
10	111.5	3.0	2910	11	055225 mus musculu
11	111	3.0	2873	12	093070 hepatitis g
12	108	2.9	3021	12	081495 hepatitis c
13	107.5	2.9	634	4	096Q25 homo sapien
14	107.5	2.9	680	4	09BVM0 O9bvm0 homo sapien
15	107.5	2.9	689	4	09BSA7 O9bsa7 homo sapien
16	107.5	2.9	733	4	096IS8 O96is8 homo sapien

17	107.5	2.9	746	4	096926 O969r6 homo sapien
18	107.5	2.9	1187	4	094812 O948r2 homo sapien
19	106.5	2.8	478	16	09R189 O9r189 deinococcus
20	106.5	2.8	1152	4	09UR1 O9ur1 homo sapien
21	106.5	2.8	1186	4	094839 O948r3 homo sapien
22	106.5	2.8	1187	4	096R23 O96r23 homo sapien
23	106	2.8	581	10	094BM5 O94bm5 cinamomum
24	105.5	2.8	1030	10	09BY59 O9by59 homo sapien
25	105.5	2.8	549	10	09FV22 O9fv22 cinamomum
26	104	2.8	1002	11	091VB4 O91vb4 mus musculu
27	103.5	2.8	758	4	09NKR9 O9nkr9 homo sapien
28	103.5	2.8	1160	10	09FMR8 O9fmr8 arabidopsis
29	103.5	2.8	3021	12	081258 O81258 hepatitis c
30	103	2.8	906	16	09HX92 O9hx92 pseudomonas
31	103	2.8	946	10	022015 O22015 cyllindrothe
32	102.5	2.7	1248	4	09UJ61 O9uj61 homo sapien
33	102	2.7	1767	5	09W4F5 O9w4f5 drosophila
34	102	2.7	1767	5	024495 O24495 drosophila
35	102	2.7	4848	2	007944 O07944 streptomyce
36	101.5	2.7	698	16	09PH5 O9ph5 xylella fas
37	101.5	2.7	1274	10	09S287 O9s287 arabidopsis
38	101	2.7	2873	12	093069 O93069 hepatitis g
39	100.5	2.7	454	5	09UA07 O9ua07 caenorhabd1
40	100.5	2.7	462	5	09Y198 O9y198 caenorhabd1
41	100.5	2.7	475	5	09U343 O9u343 caenorhabd1
42	100.5	2.7	634	5	09U9Y8 O9u9y8 caenorhabd1
43	100.5	2.7	647	5	062395 O62395 caenorhabd1
44	100.5	2.7	864	4	043297 O43297 homo sapien
45	100.5	2.7	936	13	098ST1 O98st1 gallus gall

ALIGNMENTS

RESULT 1
ID 099J43 PRELIMINARY; PRT; 567 AA.
AC 099J43;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DR 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DI 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 62.8 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strusberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004759; AAH04759.1;
KM Hypothetical protein.
SQ SEQUENCE 567 AA; 62798 MW; C1AAB79E2006B1D CRC64;

Query Match	Score	Length	DB ID	Description
Best Local Similarity	70.3%;	567		
Matches	487; Conservative	0; Mismatches	2; Indels	0; Gaps
OY	1 MPVSWFLSLALGRNPVVSLERLMEPODARCSLGSCHLMDGVLCPLGSLQSAAPGV 60			
DB	1 MPVSWFLSLALGRNPVVSLERLMEPODARCSLGSCHLMDGVLCPLGSLQSAAPGV 60			
OY	61 LVPTFRQTELVLCRPOKTDALRYRVVHLAVHGMMEPERGKSDSELRNLSLQAO 120			
DB	61 LVPTFRQTELVLCRPOKTDALRYRVVHLAVHGMMEPERGKSDSELRNLSLQAO 120			
OY	121 VVLSFOAYPIARCALLEVQPADLVOPGQSVSAVFDFEASLSLGAEOIWSYTKRYOKE 180			
DB	121 VVLSFOAYPIARCALLEVQPADLVOPGQSVSAVFDFEASLSLGAEOIWSYTKRYOKE 180			
OY	181 LNEIQQLPDCRGLEVRDSSISCVLPLNLVSTDDNVLITLDVSEQDFSLIYLRPVPD 240			


```

Db 181 LNTLQOLPDCRGLEVDSDIOSCWVLPWLVNSTDGDNVLLTLDVSEEDFSLLYLRPVD 240
QY 241 ALKSLMYKNLTGPQNTLTNTDLPVLCICIQVMSLEPDERVEFCFREDPGAHRLMHTA 300
Db 241 ALKSLMYKNLTGPQNTLTNTDLPVLCICIQVMSLEPDERVEFCFREDPGAHRLMHTA 300
QY 301 RLRLVSPGWOLADAPCCLEGGKTYLCAQAPQSPQPLVPVPQKATVNEPODQLVAGH 360
Db 301 RLRLVSPGWOLADAPCCLEGGKTYLCAQAPQSPQPLVPVPQKATVNEPODQLVAGH 360
QY 361 PNLQVQVSTWMEVQLOACSWADSLGSGKDDMLVEMKTLGNTSVSCALPSCGTPPSMA 420
Db 361 PNLQVQVSTWMEVQLOACSWADSLGSGKDDMLVEMKTLGNTSVSCALPSCGTPPSMA 420
QY 421 STRAARLCEELLQDFRSQCMQMLMNDNMGSLMACPMDKYTHRRVYLVWLACLLAALF 480
Db 421 STRAARLCEELLQDFRSQCMQMLMNDNMGSLMACPMDKYTHRRVYLVWLACLLAALF 480
QY 481 FFLLLKKDR 489
Db 481 FFLLLKKDR 489

```

RESULT 2

Q9BR97 PRELIMINARY: PRT: 538 AA.

AC Q9BR97: 01-JUN-2001 (TREMBLER. 17, Created)

DT 01-JUN-2001 (TREMBLER. 17, last sequence update)

DE 01-DEC-2001 (TREMBLER. 19, last annotation update)

DE HYPOTHEICAL 59.1 KDA PROTEIN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_Taxid=9606;

RP SEQUENCE FROM N.A.

RC TRISSUE-ENDOMETRIAL ADENOCARCINOMA;

RA Strausberg R.;

KL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL: BC006411; AAH06411.1; -

KW Hypothetical protein.

QK SEQUENCE 538 AA; 59127 MW; 65E6344DA6A5AFD2 CRC64;

Query Match 47.8%; Score 1786.5; DB 4; Length 538;
 Best Local Similarity 64.7%; Pred. No. 9.7e-154;
 Matches 346; Conservative 51; Mismatches 109; Indels 29; Gaps 5;

```

QY 1 MPVSWFLSLALGRNPVVSLERLNEPDTARCSLGLSCHLMQGVLCPLGSLQSAAPGV 60
Db 1 MPVSWFLSLALGRNPVVSLERLNEPDTARCSLGLSCHLMQGVLCPLGSLQSAAPGV 60
QY 61 LVPRLQTELVRCQKQKDCALRVVAVLAHGHAAPEE-----AGKSDELSQESRNAS 116
Db 61 LVPRLQTELVRCQKQKDCALRVVAVLAHGHAAPEE-----AGKSDELSQESRNAS 116
QY 117 LQAQVVSFOAYPIARCTALLEVQVADLVQPGOSVSAVDFCEASLGAVQVQSYTKPR 176
Db 117 LQAQVVSFOAYPIARCTALLEVQVADLVQPGOSVSAVDFCEASLGAVQVQSYTKPR 176
QY 121 LQAQVVSFOAYPIARCTALLEVQVADLVQPGOSVSAVDFCEASLGAVQVQSYTKPR 180
Db 121 LQAQVVSFOAYPIARCTALLEVQVADLVQPGOSVSAVDFCEASLGAVQVQSYTKPR 180
QY 177 YQELNLTQQLPDCRGLEVDSDIOSCWVLPWLVNSTDGDNVLLTLDVSEEDFSLLYLR 236
Db 177 YQELNLTQQLPDCRGLEVDSDIOSCWVLPWLVNSTDGDNVLLTLDVSEEDFSLLYLR 236
QY 181 YKELNLTQQLPDCRGLEVDSDIOSCWVLPWLVNSTDGDNVLLTLDVSEEDFSLLYLR 225
Db 181 YKELNLTQQLPDCRGLEVDSDIOSCWVLPWLVNSTDGDNVLLTLDVSEEDFSLLYLR 225
QY 237 PYVDALSLATKNTLTGPQNTLTNTDLPVLCICIQVMSLEPDERVEFCFREDPGAHRL 296
Db 237 PYVDALSLATKNTLTGPQNTLTNTDLPVLCICIQVMSLEPDERVEFCFREDPGAHRL 296
QY 226 QVQGPFRPMHKNLTGPQNTLTNTDLPVLCICIQVMSLEPDERVEFCFREDPGAHRL 285
Db 226 QVQGPFRPMHKNLTGPQNTLTNTDLPVLCICIQVMSLEPDERVEFCFREDPGAHRL 285
QY 297 WHIARLRLVLSFGVWOLADAPCCLEGGKTYLCAQAPQSPQPLVPVPQKATVNEPODQL 356
Db 297 WHIARLRLVLSFGVWOLADAPCCLEGGKTYLCAQAPQSPQPLVPVPQKATVNEPODQL 356
QY 286 WQARLRLVLSFGVWOLADAPCCLEGGKTYLCAQAPQSPQPLVPVPQKATVNEPODQL 345
Db 286 WQARLRLVLSFGVWOLADAPCCLEGGKTYLCAQAPQSPQPLVPVPQKATVNEPODQL 345

```

```

QY 357 VAGHNLQVQVSTWMEVQLOACSWADSLGSGKDDMLVEMKTLGNTSVSCALPSCGTP 416
Db 357 VAGHNLQVQVSTWMEVQLOACSWADSLGSGKDDMLVEMKTLGNTSVSCALPSCGTP 416
QY 346 LKGNHNLQVQVSTWMEVQLOACSWADSLGSGKDDMLVEMKTLGNTSVSCALPSCGTP 405
Db 346 LKGNHNLQVQVSTWMEVQLOACSWADSLGSGKDDMLVEMKTLGNTSVSCALPSCGTP 405
QY 417 PSMASRAARLCEELLQDFRSQCMQMLMNDNMGSLMACPMDKYTHRRVYLVWLACLLA 476
Db 417 PSMASRAARLCEELLQDFRSQCMQMLMNDNMGSLMACPMDKYTHRRVYLVWLACLLA 476
QY 406 PSKASTRAARLCEELLQDFRSQCMQMLMNDNMGSLMACPMDKYTHRRVYLVWLACLLA 464
Db 406 PSKASTRAARLCEELLQDFRSQCMQMLMNDNMGSLMACPMDKYTHRRVYLVWLACLLA 464
QY 477 AALFEFLLLKKDRRAARLCEELLQDFRSQCMQMLMNDNMGSLMACPMDKYTHRRVYLVWLACLLA 525
Db 477 AALFEFLLLKKDRRAARLCEELLQDFRSQCMQMLMNDNMGSLMACPMDKYTHRRVYLVWLACLLA 525
QY 465 AALFEFLLLKKDRRAARLCEELLQDFRSQCMQMLMNDNMGSLMACPMDKYTHRRVYLVWLACLLA 516
Db 465 AALFEFLLLKKDRRAARLCEELLQDFRSQCMQMLMNDNMGSLMACPMDKYTHRRVYLVWLACLLA 516

```

RESULT 3

Q60943 PRELIMINARY: PRT: 864 AA.

AC Q60943: 01-NOV-1996 (TREMBLER. 01, Created)

DT 01-NOV-1996 (TREMBLER. 01, last sequence update)

DE 01-DEC-2001 (TREMBLER. 19, last annotation update)

DE INTERLEUKIN 17 RECEPTOR.

GN IL17R.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_Taxid=10090;

RP SEQUENCE FROM N.A.

RC TRISSUE-THYMOMA EL4;

RX MEDLINE=96111968; PubMed=8777726;

RA Yao Z., Fanslow W.C., Seidlin M.F., Rousseau A.M., Painter S.L., Comeau M.R., Cohen J.I., Spriggs M.K.;

RT "Hepesvirus Saimiri encodes a new cytokine, IL-17, which binds to a novel cytokine receptor."

RT Immunity 3:811-821(1995).

DR EMBL: U01993; AAC52357.1; -

DR MGI:107399; 1117r.

KW Receptor.

QK SEQUENCE 864 AA; 97807 MW; 343FD51AA687DA31 CRC64;

Query Match 4.7%; Score 177; DB 11; Length 864;
 Best Local Similarity 20.5%; Pred. No. 4e-07;
 Matches 121; Conservative 65; Mismatches 189; Indels 216; Gaps 29;

```

QY 198 SIQSCW-----VLPWLVNSTDGDNVLLTLDVSEEDFSLLYLRPVP----- 239
Db 198 SIQSCW-----VLPWLVNSTDGDNVLLTLDVSEEDFSLLYLRPVP----- 239
QY 240 DALKSLWY--KNLT--GPQNTLTN-----HTDLPVLCICIQVMSLEP----- 277
Db 240 DALKSLWY--KNLT--GPQNTLTN-----HTDLPVLCICIQVMSLEP----- 277
QY 53 KNTCLDSDSHPRKNTLPPSPKNTIYNLSVSTQGEELVPLAVE--WTLOTDAISLYLEG 111
Db 53 KNTCLDSDSHPRKNTLPPSPKNTIYNLSVSTQGEELVPLAVE--WTLOTDAISLYLEG 111
QY 278 -----SER--VEFCFREDPGAHRLMHTAHLR--VLSPG-----VMQLDAPCC 318
Db 278 -----SER--VEFCFREDPGAHRLMHTAHLR--VLSPG-----VMQLDAPCC 318
QY 112 AELSVQDLNTERLCYF--QLSLQHLHRRKRRFSFHHVDPGQLEYTVHHLKPR-- 167
Db 112 AELSVQDLNTERLCYF--QLSLQHLHRRKRRFSFHHVDPGQLEYTVHHLKPR-- 167
QY 319 PGKTYLCAQAPQSPQPLVPVPQKATVNEPODQLVAGHNLQVQVSTWMEVQ 374
Db 319 PGKTYLCAQAPQSPQPLVPVPQKATVNEPODQLVAGHNLQVQVSTWMEVQ 374
QY 168 -----IPDGPVHNRKSLIFVDPDCDSKMKMTSCVSSGSLMPNITVELLDQHLR 218
Db 168 -----IPDGPVHNRKSLIFVDPDCDSKMKMTSCVSSGSLMPNITVELLDQHLR 218
QY 375 LQAQVVSFOAYPIARCTALLEVQVADLVQPGOSVSAVDFCEASLGAVQVQSYTKPR 405
Db 375 LQAQVVSFOAYPIARCTALLEVQVADLVQPGOSVSAVDFCEASLGAVQVQSYTKPR 405
QY 219 VDFTLNNEST--PYQ--VLESFSDSEHNSCFDVKQIFAPRQEEHOBANVTFLSKFH 274
Db 219 VDFTLNNEST--PYQ--VLESFSDSEHNSCFDVKQIFAPRQEEHOBANVTFLSKFH 274
QY 406 -CA-----LEP--SGC-----PPLPSMASTRARLCEELLQDFRSQCMQMLMND 447
Db 406 -CA-----LEP--SGC-----PPLPSMASTRARLCEELLQDFRSQCMQMLMND 447
QY 275 WCHHNVQVQVSTWMEVQLOACSWADSLGSGKDDMLVEMKTLGNTSVSCALPSCGTP 313
Db 275 WCHHNVQVQVSTWMEVQLOACSWADSLGSGKDDMLVEMKTLGNTSVSCALPSCGTP 313
QY 448 NMGSIMACPMDKYTHRRVYLVWLACLLAALFEFL-----LLKKDRRAARLCE 496
Db 448 NMGSIMACPMDKYTHRRVYLVWLACLLAALFEFL-----LLKKDRRAARLCE 496
QY 314 -----PVADYI--PLWVYGLITLILILVGSYVILILCMKWRSLSGADQEKHGDSKIN 364
Db 314 -----PVADYI--PLWVYGLITLILILVGSYVILILCMKWRSLSGADQEKHGDSKIN 364

```


QY 497 -----KALLHSADGAGYERLVGALASAL-SOMPLRAVADLMSRRELSA 540
DB 365 GILPVADLTPEPLRKRVKVIYADHPLIYEVLKFAOFLITACGFEVADLLEQVISE 424
QY 541 HGALAWHHORRRLIOEGVITLFS-----PAVAOCCOMLQOTVEPG 585
DB 425 VGVATWVSROKOEVENSKIIILCSRGTOAKMKAILGWAEPVQLRCDHM-----KPA 478
QY 586 PHALAAMLSCVLPDLOGATGRVGVYFDGLLHPDPSPPRVAPLESL 636
DB 479 -GDLFTAAAMNLPDPRKPCFGRVVCYFSGICSERDVDFLNTSRPL 528

RESULT 4
043844 PRELIMINARY; PRT; 866 AA.
AC 043844;
DB 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
IL-17 RECEPTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98035683; PubMed=9367539;
RA Yao Z., Spriggs M.K., Derry J.M.J., Strockbine L., Park L.S.,
Rt Vandenbos T., Zapone J., Painter S.L., Armitage R.J.;
RT "Molecular characterization of the human interleukin (IL)-17
receptor."
RL Cytokine 9:794-800(1997).
DR EMBL: U58917; AAB9730.1; --
KW Receptor.
SQ SEQUENCE 866 AA; 96122 MW; 88AF626A83F3FF70 CRC64;

Query Match 3.9%; Score 147.5; DB 4; Length 866;
Best Local Similarity 19.3%; Pred. No. 0.00019;
Matches 140; Conservative 78; Mismatches 239; Indels 269; Gaps 32;

QY 55 SAGP-----VLP-----TRLOTELVLRCPQKTDCAIRVAVVHLAVGHAAEP 100
DB 10 AVPGPLIGLILLGLVLAIPGASLRILDRALVCSQP---GINCTYKNSTCLDDSWIHPR 66
QY 101 EAGKSDSELOESRNASIQAOVVISF---QAVPIARCALLEVOVADLVQPOGVSAAVF 156
DB 67 -----NLTPSSPKDQIOIHLFAHTQOGDLEPVAN---IEWTLQTD----- 103
QY 157 DCEASTL---GAEOVQWSTKPRYOKELNLTQOLPDCRGLEVRDSIQ---SCWVLPWLN 209
DB 104 ---ASILYLEGALSV---LQINTNERL---CVREFELSKLRHHRMRFTFSH 148
QY 210 VSTDGDNVLLTLVDSEODFSFLYLRLRPVDAKSLMYKNLTGPONITLNTDVLVPCIC- 268
DB 149 FVYVDP-----QYEYTVHHLPRPIPGDPRHQSKNFLVPDCEHARMKVTPPCMS 199
QY 269 -----IOVMSLEPDSERVEFC-----PFREDPGAHRLMIARLVLS 306
DB 200 GSLMDPNITVETLEAQLRVSFTLWNESTHYQIILTSFPHMNHSCFEEMHHI----- 252
QY 307 PGWQDAPDPCILPGKVTLCQAAPDQSPQPLVPPVPOKNATVNEPODFO-----LVAGH 360
DB 253 -----PAPR-----PEEFHORSNVTLIRN 272
QY 361 PNLGVOSTWEKVQLOACSWADSLGPPKMDMLVEMKTGLNNTSVCALEPSCGTPLPSMA 420
DB 273 LAGCCR---HOVOIQ-----PF-----FSSCLND---CLHNSATVSCPEMP 307
QY 421 STRAARLAGEELLQDFRSHQCMQIANDNDMGSLMACPMDKYIHRRWLVVLAACILLAAALF 480
DB 308 DT-----PEPIDY-----MPLM-----VWFITGISILVGSVI- 337

QY 481 FFLIKKDRKRAAGS-----RTALLHSADGAGYERLVG 515
DB 338 -LIVCMFTWRLAGSGSEKYSDDTKYTDGLPAADLIPPLKPRKVIITYSADHPLVYDVL 396
QY 516 ALAS-ALSQPLRAYAVDMSRRELSAHGALAFHHORRRLIOEGVITLFSRAVAOCC 574
DB 397 KFAOFLITACGTEVADLLEBOAIS EAGVMTWVGROKOEVENSKIIYLCBSGTRAKWQ 456
QY 575 QWL-----QLOTEVEPGP-HDALAAMLSCLVLPDLOGATGRVGVYFDGLHPDPSPP 627
DB 457 ALLGRGAPVRLRCHGRKVGSLFTAAAMNLPDPRKPCFGRVVCYFSGICSDGDVFDL 516
QY 628 FRYAPLPSLPTQLEPAF-----LDALOGCGCSTSAGRPADRVERVTO-----A 668
DB 517 FGAPRPLMDRFEVEYFRIDLEMFQPG-----RMRVGLSGDNVLRSPGRQ 566
QY 669 LRSAID 674
DB 567 LRALD 572

RESULT 5
096F46 PRELIMINARY; PRT; 866 AA.
AC 096F46;
DB 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
IL-17 RECEPTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=OTERUS, AND LETOMTOSARCOMA;
RA Strausberg R.;
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: BC011624; AAH11624.1; --
KW Receptor.
SQ SEQUENCE 866 AA; 96131 MW; 28330BED2303B0C9 CRC64;

Query Match 3.9%; Score 147.5; DB 4; Length 866;
Best Local Similarity 19.3%; Pred. No. 0.00019;
Matches 140; Conservative 78; Mismatches 239; Indels 269; Gaps 32;

QY 55 SAGP-----VLP-----TRLOTELVLRCPQKTDCAIRVAVVHLAVGHAAEP 100
DB 10 AVPGPLIGLILLGLVLAIPGASLRILDRALVCSQP---GINCTYKNSTCLDDSWIHPR 66
QY 101 EAGKSDSELOESRNASIQAOVVISF---QAVPIARCALLEVOVADLVQPOGVSAAVF 156
DB 67 -----NLTPSSPKDQIOIHLFAHTQOGDLEPVAN---IEWTLQTD----- 103
QY 157 DCEASTL---GAEOVQWSTKPRYOKELNLTQOLPDCRGLEVRDSIQ---SCWVLPWLN 209
DB 104 ---ASILYLEGALSV---LQINTNERL---CVREFELSKLRHHRMRFTFSH 148
QY 210 VSTDGDNVLLTLVDSEODFSFLYLRLRPVDAKSLMYKNLTGPONITLNTDVLVPCIC- 268
DB 149 FVYVDP-----QYEYTVHHLPRPIPGDPRHQSKNFLVPDCEHARMKVTPPCMS 199
QY 269 -----IOVMSLEPDSERVEFC-----PFREDPGAHRLMIARLVLS 306
DB 200 GSLMDPNITVETLEAQLRVSFTLWNESTHYQIILTSFPHMNHSCFEEMHHI----- 252
QY 307 PGWQDAPDPCILPGKVTLCQAAPDQSPQPLVPPVPOKNATVNEPODFO-----LVAGH 360
DB 253 -----PAPR-----PEEFHORSNVTLIRN 272
QY 361 PNLGVOSTWEKVQLOACSWADSLGPPKMDMLVEMKTGLNNTSVCALEPSCGTPLPSMA 420

Db 273 LKGCRR-----HOVOIO-----PF-----FSSCLND-----CLRHSATVSCPEMP 307
 QY 421 STRAARIGEELLADFRSHOCMOAMNDNMGSLWACPMCKYHRRMVLWLAACLLAALF 480
 Db 308 DT-----PEPIPDY-----MPLM-----VWFITGISILVGSYV- 337
 QY 481 FELLKRRKRRKARS-----RVALLLHSADGAGYERLVG 515
 Db 338 -LLVTCMTWRLAGPSEKXSDTKYDGLPVADLIPPLKPRKWIITYSADHPLTVDVVL 396
 QY 516 ALAS-ALSONPLRVAVDMSBRELSTHGLAMFHHORRRLIOEGCVVILFSPAAVACO 574
 Db 397 KFAOFLTRACGTEVALDLEBOKHISEAGVMTWGRKOEKVENESKTIIVLCSRGTRAKQ 456
 QY 575 QWL-----QLOTVEBGP-HDALAAWLSGVLPDLOGRATGRVGVYFDGLHPDSVSP 627
 Db 457 ALLGRGAPVRLRCHDGKPGVGLFTTAMNMILPDKRACGCTYVVCYFSEVSCGDVFDL 516
 QY 628 PRVAPLESLPQLPAP-----LDALOGCSTSGAGRPADVERVTO-----A 668
 Db 517 FGAPRPRYELMNFEEVYRIDIEMFOPG-----RMHRYGELSGDNYLRSPPGRQ 566
 QY 669 LRSALD 674
 Db 567 LRALD 572

RESULT 6
 O53499 PRELIMINARY: PRT: 1013 AA.
 AC O53499
 DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE HELICASE.
 GN HELZ OR RV2101 OR MWV020.01.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 RX NCBI_TaxID=1773;
 RN RP
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA *Deciphering the biology of Mycobacterium tuberculosis. from the
 RA complete genome sequence.*
 RL Nature 393:537-544(1998).
 CC -1- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.
 DB EMBL: AL021924; CAA17284.1;
 DR TubercuList: RV2101;
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR00330; SMF2_N.
 DR Pfam: PF00271; Helicase_C.1.
 DR Pfam: PF00176; SMF2_N.1.
 DR SMART: SM00487; DEXDC.1.
 DR SMART: SM00490; HELIC_C.1.
 DR ATP-binding; Complete proteome; Helicase.
 KM ATP-binding; Complete proteome; Helicase.
 SQ SEQUENCE 1013 AA; 111629 MW; 601PFDLDBSCABEPI CRC64;

Query Match 3.2%; Score 119.5; DB 16; Length 1013;
 Best local Similarity 20.8%; Pred. No. 0.085;
 Matches 150; Conservative 80; Mismatches 224; Indels 267; Gaps 38;

QY 90 LAVHGHMAEP--EAGKSDSELOESRNASLOAOVLSFOAYPIARCALLVEYVPADL--- 144
 Db 2 LVHGHFWSNNGSMRLMAEDSLL-----VSPSALASARHPFA-----APADLIAG 49
 QY 145 VQPGQ-----SVGSAVDFCFE-----ASLGAEOIWSYTKRKYOKELNLT--- 184
 Db 50 IHPKPRATAVLLPRLSRAPLSDPELIRLAPRAARTDPLLAWTVF--VVDLPTALA 107
 QY 185 ---QQLPDCR-GLEY-----RDSIQSCVILFWLNVSTG-----DNVL 218
 Db 108 AFDPAPRVRYGASVDYLAELAVFAREIVERGRVILQRLRDYTGAAACMPRVLOGRVVA 167
 QY 219 LTLVSEEDQDFSEFLYLRFV-----PDALSKVMKNLTPQNTITLHNDLV 264
 Db 168 MT-----SLVSAMPVPCAEVGGHDPHELATSLDAM-----VDAVRAALSPMDL 214
 QY 265 PCLCTQVWSLEPDSERVECPREDPGAHNMHIARLRYLSPGVQDLAPCCLPKVTLL 324
 Db 215 P-----PRGRSKRRH-----AVEAW-LTALTCPDGRFDA 243
 QY 325 CMOAPDPSRCQPIVPPVPOKNATVNEPDPOLVAGHPNLCVOYSTKEVQLOACSNADSL 384
 Db 244 -----EPDELALA-----EALRPWDVGI-----GTV 266
 QY 385 GFPEKDMILVEKTKGLANTSVCALPEPGCTPLPSMASTRARIGEELLDFRSHQCM--- 441
 Db 267 GPRAATFRLSEVENTEET-----PAG-----SLMRL-EFLLOSTQDPSLVPA 309
 QY 442 -QLMNDNMGSLMACPMCKYHRRV---LVWLACILLAALFELLKRRKARGS 496
 Db 310 EQANND--GSL-----RRLMDRPOELLFELGRASRIFFELY-----PALRTA 351
 QY 497 RTALLHSADAGYERLVGALASALSQMDPLRAVVDMSBRELISAHQALMFHQRRLIO 556
 Db 352 CPSELELDADA--YRFLSGTAAYVDEAGFGVLLPSM-----W--DRRKL- 393
 QY 557 EGGVILLFSP--AAVACCOOMLQLOTVE-----PGPHDALAAMLSCVLP----- 599
 Db 394 --GLVLAATYTVDVGVGKASFRQQLVYFRRELAVAGDDPLSEETIAALTETKSPILRLR 451
 QY 600 -----DFLOGRATGRVGVYFDGLL--HPDSVSPPRVAPLESLPQLPA 642
 Db 452 GOWVALDTEQMRGLEFLERKPTGRTAELLALASHDQDVTPLEYAV-----RADG 506
 QY 643 FLDALOGCCSTSGAGRPADVERVTOALASALDSCTS-----SSAPGCCBEMDLGCTTL 697
 Db 507 WLDDLLAGAAAASIQPDPPDGFATLRPYOQGLAMTAFSLSLGSLADMDMGIGKTV 566
 QY 698 E 698
 Db 567 Q 567

RESULT 7
 O68870 PRELIMINARY: PRT: 3021 AA.
 AC O68870
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70) (NS1)]
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OC NCBI_TaxID=11103;
 RN RP
 RP SEQUENCE FROM N.A.
 RA Seelig R., Weber P., Seelig H.P., Ledger N., Botner C., Renz M.,
 RA *Hepatitis C virus type V genome isolated from a patient in Germany.*;
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.

QY 51 GSIOSAPGVLY--PTRL--OTELVLRCPQKTDALRVRVYVHLAVHGMABEAGKSD 106
 DB 125 GSISRTRAKVAGAPLRFSLQTESITAFMGDT-VLKECVIGDPMPTTHMQ-----KNO 177
 QY 107 SELQ-----ESRNASLQAOVYLVSFOAYPIARCALLLEVQVPADLYNOGQSVGSAVFDC 158
 DB 178 ODNLNIPGSR-----VVLBSGALQISR-----LQPGD--SGVYRCASRN 216
 QY 159 -FEASLGAEEVQIMWSTYKPRYQKELNLTQOLPDCRGLEVRDSIOSCV-----LPMLNV 210
 DB 217 PASTRTGNEAEVRILSDPGIAROLFLQRPNSVYAIKSKDAVLECCVSGYPRPSTWLR- 275
 QY 211 STDGNNVLTLDVSEODSFLLYLKRPVPAKLSMTKNTLGTGPONTLNHNDLVPCLCIQ 270
 DB 276 --GEEVI-----OLSKKKYSLGSGNLLISNVD----- 302
 QY 271 VMSLEPDERVEFCFREDPGAH-----RNMHTARLVLSPGVQLOAPCCLPGKV 322
 DB 303 -----DMSGITTCVYTYKKNENISASAEITVLP-PWFLNHPNML----- 340
 QY 323 TLCAQAPD-QSFCQPLVPPV-----OKNATVNEPQDPOLVAGHPNLQVOVSTWERYQLQA 377
 DB 341 -YAYESMDIEFECAYSGKVPYVNMKNGDVLPISDYFQIVGSGNLRILGVYKSDGEFYQ 399
 QY 378 CSMADSLGPFKDMILYEMKTLGNTSVYCALEPSCGTPPLSMASRAARL 427
 DB 400 CVAENAGNAGSQAOLIVKPAIPSSSILPSAPRDVVP--LVSSRFVRL 447
 RESULT 9
 QY 99B84 PRELIMINARY; PRT: 2012 AA.
 AC 09B84;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, last annotation update)
 DE CHANNEL-KINASE 2.
 GN CHAK2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN RP
 RP SEQUENCE FROM N.A.
 RC TISSUE-KIDNEY;
 RX MEDLINE=99147092; PubMed=10021370;
 RA Ryazanov A.G.; Pavur K.S.; Dorovkov M.V.;
 RT "Alpha-kinases: a new class of protein kinases with a novel catalytic
 domain";
 RL Curr. Biol. 9:R43-R45(1999).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP TISSUE-KIDNEY;
 RA Ryazanov A.G.; Pavur K.S.; Petrov A.N.; Dorovkov M.V.; Ryazanov A.G.;
 RT "Novel type of signaling molecules: protein kinases covalently linked
 to ion channels";
 RL Mol. Biol. 0:0-0(2001).
 DR EMBL: AF50881; AAK31202.1;
 DR InterPro: IPR000636; Cation_chan_non_11g.
 DR InterPro: IPR002111; Cat_channel_TripL.
 DR Pfam: PF00520; Ion_trans; 1.
 KW kinase.
 SQ SEQUENCE 2012 AA; 230493 MW; 8AD58325D8080AF CRC64;

QY 39 CHLMDGVLCLEPGLSOSAPGVLYPTRLQTELVLRCPQ-----KTDALRVRVYVHLA-- 91
 DB 1327 AHSKYGGFLVAVSNLKR-----VPPSAETVYLLSRVSDVLTATEDIQTEVYVHLTQ 1380
 QY 92 --VGHMA--EPEAKSDSELQESRNASLQAOVYLVSFOAYPIARCALLLEVQVPADLYQ 146
 DB 1381 TPVYSDMASVDEPKKEHEPIAHLLDQDKAEQVLPFLTSCPRPMTWSSPLS--QAKLMQ 1437
 QY 147 PGQSVGSAVFDCFEASLGAEEVQIMWSTYKPRYQKELNLTQOLPDCRGLEVRDSIOSCVLP 206
 DB 1438 TGGGVYNAF-----SEGDETVFS--IKRMQTCLEPSTCSDSSRSQHOQAOD----- 1486
 QY 207 WLNSTDGNNVLTLDVSEODSFLLYLKRPVPAKLSMTKNTLGTGPONTLNHNDLVPC 266
 DB 1487 ---SSLDNSTRSASQSSCSGV--PWLOP-----NTSFNINFLRRIRPPARSHS----- 1531
 QY 267 LCTQVMSLEPDERVEFCFREDPGAHRLNMLHARLVLSPGVQLOAPCCLPGKVTLGW 326
 DB 1532 ---FRFHKEKELMKIKICKIKLSGSS--ICOGAM----- 1560
 QY 327 QAPQSPQGPLVPPVQKATVNEPQDPOLVAGHPN--LCVOVSTWERYQLQA 383
 DB 1561 -----VKAKMLT-----KDRRLSKKKKKTQGLQVPIIT-----VNAQSOSDQ 1597
 QY 384 LGPFKDMILYEMKTLGNTSVYCALEPSCGTPPLSMASRAARLGEELIQ--DE--RSHQ 439
 DB 1598 LNPEGENSISEEYSKMMFTYKSFHTGEVPIHQ--KMKTKELGCAQISIDLKSHQ 1656
 QY 440 CM--QLMND-----DNMGLMACPMQYIHRWYLVMLACILLAAAL 479
 DB 1657 DLKSNLSMNSSTNLNRSLKSSIGVDKISASLKSPOEHHYSAL----- 1703
 QY 480 FFFLLKKDRKRAKRSRALLHSADGAGYERLVGALASALQMPRAVADLSRRRLS 539
 DB 1704 -----ERNMRLSOTIPTPTPOVLFAGEITVYRLSES--SPLNDKSMSSWSQGRA 1754
 QY 540 A-----HGAL-----ANFHQRRRIQDEGV-VILLFSPAAYACQOQMLQ 580
 DB 1755 AMIQVLSREENDGGLRKAMRVYSTSEDD--ILKGGQFIVASFLPEVYRTHIKIQES 1811
 QY 581 TVEPGPHDALAMLSGVLPDLQGRATGRYGVYEDGLLHPDSVPSPFRVAPLF 634
 DB 1812 TV-----LHLCIREIQOQRAAOKL--ITYENQVAPQITIPYTPRLEVF 1852
 RESULT 10
 QY 055225 PRELIMINARY; PRT: 2910 AA.
 AC 055225;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE OTOGELIN.
 GN OTOG.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN RP
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE=98070772; PubMed=9405633;
 RA Cohen-Salmon M.; El-Amraoui A.; Leibovici M.; Petit C.;
 RT "Otogelin: A glycoprotein specific to the acellular membranes of the
 inner ear";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:14450-14455(1997).
 DR EMBL: U96411; AAB96561.1;
 DR HSSP: P56682; ICCV.
 DR MGD: MGI:1202064; OTOG.
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR002919; TIL.

QY 1 MPVSWFLTSLALGRNPVYVSLERMEPOPTAR-----CSLGLS 38
 DB 1268 MPSS-LTRSLAGRHPRPVVQSGALLLEITNSKREATVVRNDQEOETOSIVSGVSPNRQ 1326

Query Match 3.0%; Score 111.5; DB 4; Length 2012;
 Best Local Similarity 19.9%; Pred. No. 1.2; Mismatches 276; Indels 209; Gaps 31;

Db 2590 SNGCTCYIKVACADRVGLKNSLLIAGDDCLICERPICDPSSEALGRALAIYVACEPS 2649
 QY 688 ---EMWLCG-CYT 696
 Db 2650 YHASLDIAPYCTT 2662

RESULT 12

081495 ID 081495 PRELIMINARY: PRT: 3021 AA.
 AC 081495: \$
 DE 01-NOV-1996 (TEMBLrel. 01, Created)
 DE 01-NOV-1996 (TEMBLrel. 01, last sequence update)
 DE 01-DEC-2001 (TEMBLrel. 19, last annotation update)
 DE GENOME POLYPROTEIN (CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70) (NS1)).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OK NCBI_Taxid=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K3A;
 RA Date T;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K3A;
 RA MEDLINE=95053917; PubMed-7964640;
 RA Yamada N., Maitahara K., Mizokami M., Ohba K., Takada A., Tsutsumi M.,
 Date T.;
 RE "Full-length sequence of the genome of hepatitis C virus type 3a:
 RE comparative study with different genotypes.";
 RE J. Gen. Virol. 75:3279-3284(1994).
 RE -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
 DR EMBL: D28917; BAA06044.1;
 DR HSP: P27958; IHEI.
 DR MEROPS: S29.001;
 DR MEROPS: 039.001;
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR002522; HCV_capsid.
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV_env.
 DR InterPro: IPR002531; HCV_NSI.
 DR InterPro: IPR004109; HCV_NS2.
 DR InterPro: IPR000745; HCV_NS3.
 DR InterPro: IPR001490; HCV_NS4a.
 DR InterPro: IPR002868; HCV_NS4b.
 DR InterPro: IPR002166; HCV_NS5a.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF01543; HCV_capsid.1.
 DR Pfam: PF01542; HCV_core.1.
 DR Pfam: PF01539; HCV_env.1.
 DR Pfam: PF01560; HCV_NSI.1.
 DR Pfam: PF01538; HCV_NS2.1.
 DR Pfam: PF02907; HCV_NS3.1.
 DR Pfam: PF01006; HCV_NS4a.1.
 DR Pfam: PF01001; HCV_NS4b.1.
 DR Pfam: PF01506; HCV_NS5a.1.
 DR Pfam: PF00998; HCV_NS5a.1.
 DR Pfam: PF00271; Helicase_C.1.
 DR Pfam: PF00271; Helicase_C.1.
 DR ProDom: PD186062; HCV_NSI.1.
 DR AP: binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
 DR Nonstructural protein; Polypeptide; RNA-directed RNA polymerase;
 DR Transmembrane.
 RN SEQUENCE 3021 AA: 328386 MW: A97418FF36C062A4 CRC64;

Query Match 2.9%; Score 108; DB 12; Length 3021;
 Best Local Similarity 20.1%; Pred. No. 4.3;
 Matches 144; Conservative 77; Mismatches 247; Indels 248; Gaps 36;

QY 30 TARGSLGSLCHL-----WDGVLCLPGLSQAAPVLPVTRLOTELVLKCPQKTDALR 83
 Db 303 TNCSTL-YFGHLSCGRMAWD-----MMNMSPAVGVAH-----ILRLPQTLFDVVA 349
 QY 84 -----VAVVYHLAVHGHAEPE-----EAGKSDSELDGSRNMSIAQVYLSFOAY 128
 Db 350 GAHWGIIAGLAYSMQGNNAKVAIIMVMSGVADSTHTVAGQAARN-----AY 397
 QY 129 PIARCALLLEVQPADLVQPGQS--VGSAAVDFCEA-SLSAEVQIWSYTKPRYQKELNLT- 184
 Db 398 GITSLSVGAQKQNLQINLNGSNHINRLALNCNESITGTGIAAGFYHK-----FNSTG 451
 QY 185 --QQLPDRGLEV-----RDSIOSCV-----VLPWLNVN-----T 212
 Db 452 CPGRLSSCKPIFFKQMGPLFDANITGPBDDKPYCHVAAPRCGIYALNVCGFVYCGT 511
 QY 213 DGDNVLTLFDV-----SEEDFSFLIYLRVPALMSIMYKNLTG-----PON 255
 Db 512 PSPVVGTTDAKGAPLYTMGANKTDVFLLESRLRPSGRMFCGTMMNSTGYTKTGAPPCN 571
 QY 256 I-----TLNHTDVLVPCLIQVWSLEPDSERVECP--FREDGARNHMLHIALRYLS 306
 Db 572 IYGDGRDAQNESDL-----FCPTDCFRKHPEATYSR-----CG 604
 QY 307 PGVQQLDAPCCLPKGYTLICWQAPDQSPQPLVPVPQKNATVNEPQDFQLYAGHPNLCVQ 366
 Db 605 AGPWL--TPRCVLDVYPRRLMHP-----CTVN-----FTLE--KVMRF 638
 QY 367 VSTNEKYVQACSMADSLPFRKDMILVEKGTGLNNTSVCLSPSGCPPLSMASRAR 426
 Db 639 VGGFHRFLTAACNWTGEGKEDIEDRDSRSEPHLHSTTELAIIICSTFPMAL-STGLIH 697
 QY 427 LGSELLQDFSRHQCMQMDNDNGSL-WACPMKYIHRRWLVMLACLLAA--LEFF 482
 Db 698 LHQNLVD-----VQLVGISGGMGVAL-----KMEFVILIFELLADARCVALLM 742
 QY 483 LILKDRKRAAGSRALLIHSADG-----AGER-----LVGALAS 519
 Db 743 LILITSOAEALLENVTLNAAVAAAGTHIGWYLVAFCAAMVVRKLYPLVYTSILGLWSL 802
 QY 520 ALSQPLRAVAVDMSRREISAAGA-----LAMFHQRRIILQEGVYILFESPA 568
 Db 803 ALLVLLLPORAVAMSGEDSATLCAGLIYLFGEFTLSPTWKYIAR-----LIMNOY 854
 QY 569 AVAQCCQWQLQTVPEGPDALAAMLSCVLPDEFQGRATGRVGVYFEDGLHPDSV 624
 Db 855 TICRCE-----SALHVM--VPILARG--GNDGVILLSLSLYPSLI 891

RESULT 13

096025 ID 096025 PRELIMINARY: PRT: 634 AA.
 AC 096025:
 DE 01-DEC-2001 (TEMBLrel. 19, Created)
 DE 01-DEC-2001 (TEMBLrel. 19, last sequence update)
 DE 01-DEC-2001 (TEMBLrel. 19, last annotation update)
 DE KM-PA-2 PROTEIN.
 GN KM-PA-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nakatsura T., Senju S., Shichijo S., Ito M., Ohkouchi S., Yamada K.,
 Ogawa M., Itoh K., Nishimura Y.;
 RT "Cellular Identification of Human Pancreatic Cancer Antigen KM-PA-2,
 RT Originally Defined by SEREX Method.";
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=21134339; PubMed-11237751;
 RA Nakatsura T., Senju S., Yamada K., Jotsuka T., Ogawa M., Nishimura Y.;

Gene cloning of immunogenic antigens overexpressed in pancreatic cancer.
 RT Biochem. Biophys. Res. Commun. 281:936-944 (2001).
 DR EMBL: AB060694; BAB70666.1;
 SO SEQUENCE 634 AA; 71930 MW; E7A5608B28D65C36 CRC64;

Query Match 2.9%; Score 107.5; DB 4; Length 634;
 Best Local Similarity 21.1%; Pred. No. 0.55;
 Matches 129; Conservative 72; Mismatches 232; Indels 177; Gaps 33;

42 WDGVLCPSGLSAGVLPVTRP...
 115 FSGDVMIHPVNNRPADKRSFISLVEKEKYSR...
 102 AGKS...
 161 PROTPSFYDLMAGEDPNAVIGRH...
 156 FDCE...
 200 -ESYNPPEYLLSEERLAMEQOEGEKLSFLPKFPSLRAPVAGRTQERFERCIDL 258
 203 WVL...
 259 YLCPROKRMVNDPELIPKIPRPD...
 260 HTDLPVLCIOWMSLEPDSERVECPREDPGAHNLHIARL...
 300 HSDLVKCL...
 311 QLDAPCLPEKVT...
 351 NPSPAVCIVAAVEDSVLLNPAIGDRIVAGSTDDLSAFVPEEPPLQPARMLEASEE 410
 351 PQ-DPOLVAGHPNLCVOVSTWE...
 411 ROVGLRLRICHGKRPVQV...
 470 VGR...
 451 SLWACPM-DKTIHRRV...
 513 SLAVHPAGDNVIGSGYSDSKVWFDLDLSTKPYRLRHKKALRAVAFHRYPLFASGSD 572
 508 AGYERLVGALASLSQMP...
 573 GSVIVCHGVNVDLQNP...
 568 AAVACQOQWL 577
 617 T-----QPMV 621

RESULT 14
 ID 09BVM0 PRELIMINARY: PRT: 680 AA.
 AC 09BVM0
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE SIMILAR TO BLOCK OF PROLIFERATION 1 (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RA SEQUENCE FROM N.A.
 RC TISSUE=MUSCLE, RHABDOMYOSARCOMA;
 RA Strausberg R.;
 RI Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 DR EMBL: BC001086; AA01086.1;
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 7.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.
 DR PROSITE: PS50082; WD_REPEATS_2; 2.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 2.
 KW Repeat; WD repeat.
 FT NON_TER
 SO SEQUENCE 680 AA; 76791 MW; A26B8F6FCB1F2F CRC64;

Query Match 2.9%; Score 107.5; DB 4; Length 680;
 Best Local Similarity 21.1%; Pred. No. 0.6;
 Matches 129; Conservative 72; Mismatches 232; Indels 177; Gaps 33;

42 WDGVLCPSGLSAGVLPVTRP...
 161 FSGDVMIHPVNNRPADKRSFISLVEKEKYSR...
 102 AGKS...
 207 PROTPSFYDLMAGEDPNAVIGRH...
 156 FDCE...
 246 -ESYNPPEYLLSEERLAMEQOEGEKLSFLPKFPSLRAPVAGRTQERFERCIDL 304
 203 WVL...
 305 YLCPROKRMVNDPELIPKIPRPD...
 260 HTDLPVLCIOWMSLEPDSERVECPREDPGAHNLHIARL...
 346 HSDLVKCL...
 311 QLDAPCLPEKVT...
 397 NPSPAVCIVAAVEDSVLLNPAIGDRIVAGSTDDLSAFVPEEPPLQPARMLEASEE 456
 351 PQ-DPOLVAGHPNLCVOVSTWE...
 457 ROVGLRLRICHGKRPVQV...
 516 VGR...
 451 SLWACPM-DKTIHRRV...
 559 SLAVHPAGDNVIGSGYSDSKVWFDLDLSTKPYRLRHKKALRAVAFHRYPLFASGSD 618
 508 AGYERLVGALASLSQMP...
 619 GSVIVCHGVNVDLQNP...
 568 AAVACQOQWL 577
 663 T-----QPMV 667

RESULT 15
 ID 09BSA7 PRELIMINARY: PRT: 689 AA.
 AC 09BSA7
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE SIMILAR TO BLOCK OF PROLIFERATION 1 (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606;

SEQUENCE FROM N.A.

TISSUE=BRAIN, NEUROBLASTOMA;

Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

-1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

EMBL: BC005160; AA05160.1; 1-2

InterPro: IPR001680; WD40.

SMART: SM00320; WD40; 7.

PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.

PROSITE: PS0082; WD_REPEATS_2; 2.

PROSITE: PS0294; WD_REPEATS_REGION; 2.

Repeat: WD repeat.

NON_TER 1

SEQUENCE 689 AA: 77638 MW: 8066FFDEE623B7B CRC64:

Query Match 2.9%; Score 107.5; DB 4; Length 689;

Best Local Similarity 21.1%; Pred. No. 0.61;

Matches 129; Conservative 72; Mismatches 232; Indels 177; Gaps 33;

42 WDGDVCLPGSLQAPGVLPTRIQTELVLCPQKDCALNVRVYVLAHGHMAPEE 101

170 FSGDVMIHPYTNRPADKRSFIPSLVEKEKYSR-----MVH-AIKMGWIDPRR 215

102 AGKS-----DSELDSESNASLQAOVYLSFOAYPIARCALLEVQVPA-DLVQPGQSVGSAAV 155

216 PRDPTSFYDLMAQEDENNAVAGRH-----KMHVPAKIALPGHA----- 254

156 FDGE-----EASLGAEOI-WSTYKPRYOKEINLTQOLPDCR-----GLEVRDSIOSC--- 202

255 -ESYNPPETLLSEERIAMEQOEGERKLSFLPKRFSIRAVPAIGRFIOERERCLDL 313

203 WVLPP--WLVNSTGDVNLTLTDVSEEDFSFLYLKRPVPDAKLSIMYKNLTGPONITLN 259

314 YLCPRQRKMRNVNVPEDILPKLRPRD-----LQPP-TCQALVYR-----G 354

260 HTDLVPCLCIQVWSLEPDSERVECPREDPCAHNMIMHARLR-----VLSPCVW 310

355 HSDLVRLCL-----SVSPGCG--QNLVSGSDDSLR-LMEVATACVTVTPVGGVKKSVAM 405

311 QLDAPCLPGKVT-----LCMQAPDQ-----SPCQPLVPVPQKNATVNE 350

406 NPSPAVCLVAAAVEDSVLLNLPALGDRLVAGSTDQLLSAFVPEEPFLQPARMLEASREE 465

351 PQ-DFQIVAGHPNLCVQVSTWE-----KVQLQACSNWDSIGPFRKDMIL 393

466 RQVGLRLRLICGKRPYQV-TWISGRGDIYLAIVLATQGHQVLIHQLSRRRSQSPFRSHQ 524

394 VEMKTGLNNTSVCALEPSCGTPPLPSNASTRARLGEELQDFNSH---OCQQLMNDNMKG 450

525 VQR-----VAFHPA--RPEFLVAAQSQRVRLYLROELTFKIMPNC--W-----VS 567

451 SLWACPM-DKTYHRRM--VLWVLACILLAAALFFELLKKDRKAARSGFTALLHSADG 507

568 SLAVHPADNVTICGSDSKLWDFDLSTRPYRMLRHKKRALRAVAHPRYPLEPFGSGSD 627

508 AGYERLVGALASLSOMPLRVAVDLMSRELISAHGALAMFHQRRLIQEGGVVILLFSP 567

628 GSVIVCHGKVTYNDLQNLPLVY-----KVLKGH-----VLTARDLGVLDTVFHP 671

568 AAVAQCOQWL 577

672 T-----QPMV 676

Search completed: September 28, 2002, 19:42:59

Job time: 1012 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2002, 19:43:34 : Search time 24.51 Seconds

(without alignments)
1102.662 Million cell updates/sec

Title: US-09-899-471-5
Perfect score: 3741
Sequence: 1 MPVSMFLLSLGRPVVYS.....SSFAFGCCFMDLPCTTLE 698

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114.5	3.1	1447	1 DCC_MOUSE	P70211 mus musculu
2	112	3.0	1447	1 DCC_HUMAN	P43146 homo sapien
3	109.5	2.9	1711	1 PTPO_RAT	O64612 rattus norv
4	108.5	2.9	682	1 Y12_HUMAN	Q14137 homo sapien
5	106.5	2.8	1409	1 AEX3-CAEL	O02626 caenorhabdi
6	103	2.8	557	1 OM6_CHIPS	P23701 chlamydia p
7	100.5	2.7	1248	1 APAF_HUMAN	O14727 homo sapien
8	99.5	2.7	1451	1 A2M2_MOUSE	P28666 mus musculu
9	99	2.6	2437	1 NOTC_BRARE	P45530 brachydanio
10	98.5	2.6	669	1 GLSK_HUMAN	O94925 homo sapien
11	97.5	2.6	1618	1 NEST_HUMAN	P46881 homo sapien
12	96.5	2.6	1249	1 APAF_RAT	O98955 rattus norv
13	93.5	2.5	204	1 C267_HUMAN	O98955 rattus norv
14	93.5	2.5	636	1 PABP_HUMAN	O98955 rattus norv
15	93.5	2.5	726	1 CCT1_HUMAN	O60563 homo sapien
16	93.5	2.5	880	1 TYO3_MOUSE	O60563 homo sapien
17	93.5	2.5	1062	1 NAT2_MOUSE	P55144 mus musculu
18	93.5	2.5	2647	1 ABP2_HUMAN	O98955 rattus norv
19	93	2.5	443	1 W70T_RAT	P21333 homo sapien
20	93	2.5	777	1 BISC_ECOLI	O35828 rattus norv
21	92.5	2.5	566	1 TS13_MOUSE	P20099 escherichia
22	92.5	2.5	890	1 TYO3_HUMAN	O01755 mus musculu
23	92.5	2.5	3707	1 GSPM_MOUSE	O06418 homo sapien
24	92	2.5	398	1 GSPM_MOUSE	O05793 mus musculu
25	91.5	2.4	235	1 FL3JL_HUMAN	P15771 klebsiella
26	91.5	2.4	1476	1 A2M1_MOUSE	P49771 homo sapien
27	91.5	2.4	362	1 CKRA_HUMAN	P28665 mus musculu
28	91	2.4	1058	1 GEM4_HUMAN	P46092 homo sapien
29	91	2.4	1216	1 AEGP_RAT	P57678 homo sapien
30	90.5	2.4	385	1 YEH1_ECOLI	O63191 rattus norv
31	90.5	2.4	458	1 MUCM_RABIT	P33361 escherichia
32	90.5	2.4	479	1 MUCM_RABIT	P03988 oryctolagus
33	90.5	2.4	844	1 YDZ5_SCHPO	P04221 oryctolagus
					O13712 schizosacch

ALIGNMENTS

34	90.5	2.4	880	1 TYO3_RAT	P55146 rattus norv
35	90	2.4	1928	1 LPH_RAT	O02401 rattus norv
36	89.5	2.4	837	1 GCSR_MOUSE	P40223 mus musculu
37	89.5	2.4	1370	1 Z261_HUMAN	O14202 homo sapien
38	89.5	2.4	1927	1 LPH_HUMAN	P09848 homo sapien
39	89.5	2.4	4543	1 LRPL_CHICK	P98157 gallus gall
40	89	2.4	922	1 W70T_MOUSE	O942V7 mus musculu
41	89	2.4	942	1 MAKE_MOUSE	O942V7 mus musculu
42	89	2.4	1189	1 HAIR_HUMAN	O942V7 mus musculu
43	88.5	2.4	938	1 EBN4_EBV	O43593 homo sapien
44	88.5	2.4	1208	1 RCO4_HUMAN	P03203 epstein-bar
45	88.5	2.4	1258	1 YS00_ANNSP	O94761 homo sapien
					O87C2 anabaena sp

RESULT 1
ID DCC_MOUSE STANDARD: PRT: 1447 AA.
AC P70211
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Tumor suppressor protein DCC precursor.
GN DCC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-Brain;
RX MEDLINE=96112625; PubMed=8570174;
RA Cooper H.M., Ames P., Britto J., Gad J., Wilks A.F.;
RT "Cloning of the mouse homologue of the deleted in colorectal cancer
RT gene (mDCC) and its expression in the developing mouse embryo.";
RT Oncogene 11:2243-2254(1995).
RN (2)
RP REVISIONS.
RC STRAIN-BALB/C; TISSUE-Brain;
RA Cooper H.M.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS OF THE PROTEIN ARE PRODUCED FROM
CC THE SAME GENE BY THE USE OF ALTERNATIVE INITIATION SITES. A THIRD
CC FORM WHICH IS EXPRESSED ONLY IN THE EMBRYO IS PRODUCED BY
CC ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED AT HIGH LEVELS IN THE
CC DEVELOPING BRAIN AND NEURAL TUBE. IN ADULT, HIGHLY EXPRESSED IN
CC BRAIN WITH VERY LOW LEVELS FOUND IN TESTIS, HEART AND THYMUS.
CC -1- DEVELOPMENTAL STAGE: LOW LEVELS IN EARLY GESTATION. HIGHEST LEVELS
CC EXPRESSED DURING MID GESTATION. LEVELS DECREASE IN LATE GESTATION
CC AND REMAIN AT THIS LEVEL IN THE ADULT.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X85788; CAA59786.1; -
CC HSSP: P56276; 1TLK.
CC MGD: MGI:94869; DCC.
CC InterPro: IPR003961; FN_III.
CC InterPro: IPR003962; FNIII_repeat.
DR

	FTT	CONFLICT	138	138	MISSING (IN REF. 3).
FTT	CONFLICT	233	329	MISSING (IN REF. 3).	
FTT	CONFLICT	421	421	MISSING (IN REF. 3).	
SO	SEQUENCE	1447	AA; 158456	MM; 4A8612766ED0471F	CRC64;

Query Match	3.0%	Score 112;	DB 1;	Length 1447;
Best Local Similarity	18.8%	Pred. NO. 0.71;		
Matches 136;	Conservative 74;	Mismatches 210;	Indels 302;	Gaps 34;

SO SEQUENCE 1447 AA; 158456 MW; 4AA8612766ED0471F CRC64;

Matches 136; Conservative 74; Mismatches 210; Indels 302; Gaps 34;

QY 20 SLERLMEPQDTARSLGSLCHMDGDVLCPLGSLOSAPG-PVLVPTRKOTELVL---RC 74
 :| | |
 Db ALRFLSEEDSAVF-----MGSGNLT-ITCSAEEPCDINUTIRIVKPDCTATIGSNT 80
 39 :| | | :| | | :| | |

```

0Y      75 POKTDCALRRVVVHLAVHGMAEPE-----EACKSDSELQESRNAST-----Q 118
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

```

QY 119 AQVLSFQAVFI-ARCALLEVOYPA-----DL-----YOPG 148

140 1ESVIAFMGVLNCEVIGEPMTIHWQKQDDLP1PGDSRWVLPSCADQISRLQPG 205

Db 206 -DIG--IYRC SARNPASSRTGNEAEVRILSDPGLHRQLYFLQRPSPNVVAIEGKDAVLECC 262

```

db      263 VSGYPPPSFTWI.R-----GFEVT-----OI DSEVVEI ICCSNTY 306
      | |      | |      | : | :      | : | | | | :

```

```

qy      25/ TLNHTDLVPCLCIQWMSLEPDSERVEFCPFREDPGAH-----RNLWHIARLRVLSPG   308
          |||               :|||

```

QY 309 VWQLDAPCCLPGKVTLCWQAPD-QSPCQPLVPPV-----QKNATVNEPQDFQLVAGHPNL 363

364 CWOVSTWETVOT ODOCVADOT GDDVDD 383

Db 386 RI-----LGVKSDGEFYQCAENEAQNAQTSAQLIVPKPAIPSSSV 427

```

Db      428 LPSAPRDVVPVLVSSRFYLS-----WRPPAE----- 454
          |||::|||

```

400 VMTLC DULFARLF E E LULNN DRNKAARGS - R KALLHLSADGAGYERHVGALASALSQML 526

QY 527 R-----VAVDMSRRELSAHGALAEFHQRRILOEGVILLFSPAANAQO 574

00 575 0W10-----10TVEBBDHDA1 AAW SCW DNET OCB- 3MCOUICU1000 010

D6 523 PELQVPGPVENLQAVSTSPSTILITWEP----PAYANGPVQGYRFLCTEYVSTGKEQNIYD 579

Db	580	GL	581
----	-----	----	-----

TEST NO.	TEST NAME	RESULT	3
1	1.0	1.0	1.0
2	2.0	2.0	2.0
3	3.0	3.0	3.0
4	4.0	4.0	4.0
5	5.0	5.0	5.0
6	6.0	6.0	6.0
7	7.0	7.0	7.0
8	8.0	8.0	8.0
9	9.0	9.0	9.0
10	10.0	10.0	10.0
11	11.0	11.0	11.0
12	12.0	12.0	12.0
13	13.0	13.0	13.0
14	14.0	14.0	14.0
15	15.0	15.0	15.0
16	16.0	16.0	16.0
17	17.0	17.0	17.0
18	18.0	18.0	18.0
19	19.0	19.0	19.0
20	20.0	20.0	20.0
21	21.0	21.0	21.0
22	22.0	22.0	22.0
23	23.0	23.0	23.0
24	24.0	24.0	24.0
25	25.0	25.0	25.0
26	26.0	26.0	26.0
27	27.0	27.0	27.0
28	28.0	28.0	28.0
29	29.0	29.0	29.0
30	30.0	30.0	30.0
31	31.0	31.0	31.0
32	32.0	32.0	32.0
33	33.0	33.0	33.0
34	34.0	34.0	34.0
35	35.0	35.0	35.0
36	36.0	36.0	36.0
37	37.0	37.0	37.0
38	38.0	38.0	38.0
39	39.0	39.0	39.0
40	40.0	40.0	40.0
41	41.0	41.0	41.0
42	42.0	42.0	42.0
43	43.0	43.0	43.0
44	44.0	44.0	44.0
45	45.0	45.0	45.0
46	46.0	46.0	46.0
47	47.0	47.0	47.0
48	48.0	48.0	48.0
49	49.0	49.0	49.0
50	50.0	50.0	50.0
51	51.0	51.0	51.0
52	52.0	52.0	52.0
53	53.0	53.0	53.0
54	54.0	54.0	54.0
55	55.0	55.0	55.0
56	56.0	56.0	56.0
57	57.0	57.0	57.0
58	58.0	58.0	58.0
59	59.0	59.0	59.0
60	60.0	60.0	60.0
61	61.0	61.0	61.0
62	62.0	62.0	62.0
63	63.0	63.0	63.0
64	64.0	64.0	64.0
65	65.0	65.0	65.0
66	66.0	66.0	66.0
67	67.0	67.0	67.0
68	68.0	68.0	68.0
69	69.0	69.0	69.0
70	70.0	70.0	70.0
71	71.0	71.0	71.0
72	72.0	72.0	72.0
73	73.0	73.0	73.0
74	74.0	74.0	74.0
75	75.0	75.0	75.0
76	76.0	76.0	76.0
77	77.0	77.0	77.0
78	78.0	78.0	78.0
79	79.0	79.0	79.0
80	80.0	80.0	80.0
81	81.0	81.0	81.0
82	82.0	82.0	82.0
83	83.0	83.0	83.0
84	84.0	84.0	84.0
85	85.0	85.0	85.0
86	86.0	86.0	86.0
87	87.0	87.0	87.0
88	88.0	88.0	88.0
89	89.0	89.0	89.0
90	90.0	90.0	90.0
91	91.0	91.0	

AC	064612;	
PT	01-NOV-1997 (Rel	35 Created)

DE 15-JUL-1999 (Rel. 38, last annotation update)
osteosticulaar protein tyrosine phosphatase precursor (EC 3.1.3.48)

QY 547 FH0RRRILOEGGVILLFSPAANVACQOMIQLQTVBGP-HDAL-----AAMLS----- 595
 DB 812 SVRCAGAPQLQASHVLL-----SVEPGPVEDVLCHEPATYALAMWMP 855
 QY 556 -----CVLPDEFLQGNATGRNV-----GYVF-----DGLHPDSVP-SPEFVAPLFSLP 637
 DB 856 AGDVDCV-----VVELVYPCGGGTHFYQVNTSGDALLPLNLPPTSTYLS--LTVL 906
 QY 638 TQLPAPLALQGGCSTSA 655
 DB 907 GNSRMSRAVSLVCSSTA 924

RESULT 4

ID Y124_HUMAN STANDARD; PRT; 682 AA.

AC Q14137;
 DT 01-NOV-1997 (Rel. 35, Created).
 DT 01-NOV-1997 (Rel. 35, Last sequence update).
 DT 16-OCT-2001 (Rel. 40, Last annotation update).
 DE Hypothetical protein KIAA0124 (Fragment).
 GN KIAA0124.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID:9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE-96127530; PubMed-8590280;
 RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
 RT "Prediction of the coding sequences of unidentified human genes. IV.
 RT The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
 RT analysis of cDNA clones from human cell line KG-1.";
 RC DNA Res. 2:167-174(1995).
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC -1- SIMILARITY: STRONG, TO YEAST YMR049C.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: D50914; BAA09473.1;
 CC InterPro: IPR001680; WD40.
 CC Pfam: PF00400; WD40; 7.
 CC SMART: SM00320; WD40; 5.
 CC PROSITE: PS00678; WD_REPEATS_1; 1.
 CC PROSITE: PS50082; WD_REPEATS_2; 2.
 CC PROSITE: PS50294; WD_REPEATS_REGION; 2.
 CC Hypothetical protein; Repeat; WD repeat.
 KW NON_TER 1
 FT REPEAT 347 386 WD 1.
 FT REPEAT 388 428 WD 2.
 FT REPEAT 468 512 WD 3.
 FT REPEAT 513 551 WD 4.
 FT REPEAT 554 593 WD 5.
 FT REPEAT 597 636 WD 6.
 FT REPEAT 652 681 WD 7.
 SO SEQUENCE 682 AA; 76949 MW; 929219E600E732BD CRC64;

Query Match 2.9%; Score 108.5; DB 1: Length 682;
 Best local Similarity 21.1%; Pred. No. 0.5; Mismatches 232; Indels 177; Gaps 33;
 Matches 129; Conservative 72;

QY 42 WDCVYLCPGSLQSPGVLPVTRQLQTELVLRCPORTCALRVVVVLAHVHWAEPPEE 101
 DB 163 FSGVYMHPTNRPADKRSFIPSLVEKEKYSR-----MVH-AIKMKMIQPRR 208
 QY 102 AGKS-----DESLQSRNASLDAQVYVLSQAFIARCALLLEVQPA-DLVQPGSGVGSAS 155

DB 209 PRDPTPEFYDLMAQEDPNNAVIGRH-----KMHVAPRLALCGHA----- 247
 QY 156 FDCF-----EASLGAEOI-WSTPKRYQKELNTQOLPDCR-----GLEVRDSIOSC--- 202
 DB 248 -RSYNPEPPYLLSEERLAMEQOEPGERKLSFLPRKPSLRAVPAYGPFIOERFERCLDL 306
 QY 203 WYLP---NLNVSPDGNVLLTLTVSEEDFSFLYLVRPPLAKSLMKNLKLPQNTITLN 259
 DB 307 YLCPROKMRKRVNVPEDILIKLPRPD-----LOFPF-TCOLVTR-----G 347
 QY 260 HTDLVPCICIQWVSLPEPSERVEFCPPREDGAGRNLMIARLR-----VLSPGW 310
 DB 348 HSDLVKCL-----SVSPGC---QMLVSGSDGSIR-LMEVATARCVRVPVGGVKSVM 398
 QY 311 OLDAPCCLPCKYT-----LCWAPDQ-----SPCQPLVPVPQKNATVNE 350
 DB 399 NESPACVLAAAVEDSVLLNLPALGRLVAGSTQLLSAFVPPEEPPLQPARMLASSEE 458
 QY 351 PQ-DEOLVAGHNPICVQVSTWE-----KYQLQACSMADSLGPFEDMILL 393
 DB 459 KQVGLRLRICHGKRPVTVQ-TWGRGDYLAVALTOGHVYLIHQLSRRRSSQPFRRSHQO 517
 QY 394 VEMKTGLNNTSVCALEPSCCTPLPSMASTRARLGEELLQDFRSH---QCMLWMDNMG 450
 DB 518 VQR-----VAFHBA--RPLVAVASQSRVRLYHLRQELTKIMPNCK--W-----VS 560
 QY 451 SIMACPM-DKTYHRRR--VLVYLACILLAALFFELLKDKRKAARSGRALLHSADG 507
 DB 561 SLAVHPADNVTGSGVSKLWFDDLSTKPYRLRHKKKALRAVAHPRPLFASGSD 620
 QY 508 AGYERLVGALASLSQMPRAVADLMSRRELSAAGALAMFH0RRRILOEGGVILLFSP 567
 DB 621 GSVYVCHGVNVDLQNPILVTV-----KVLKGH-----VLTIRGLVLDVIFHP 664
 QY 568 AAVACQGM 577
 DB 665 T-----QPVW 669

RESULT 5

AEX3_CAEBL STANDARD; PRT; 1409 AA.

ID AEX3_CAEBL 002626; 027467;
 AC 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Regulator of presynaptic activity aex-3.
 GN AEX-3 OR C02H7.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditioidea;
 OC Rhabditiidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID:6239;
 RN [1]

RP SEQUENCE FROM N.A.
 RP STRAIN-BRISTOL N2;
 RC MEDLINE-97282461; PubMed-9136770;
 RX Iwasaki K., Staunton J., Saifee O., Nonet M., Thomas J.H.;
 RT "aex-3 encodes a novel regulator of presynaptic activity in C.
 RT elegans.";
 RL Neuron 16:613-622(1997).
 RN [2]

RP SEQUENCE FROM N.A.
 RP STRAIN-BRISTOL N2;
 RA Keimig D., Mink M.;
 RL Submitted (FEB-1996) to the EMBL/Genbank/DBJ databases.
 RN [3]

RP INTERACTION WITH CAB-1.
 RX MEDLINE-20428446; PubMed-10970871;
 RA Iwasaki K., Toyonaga R.;
 RT "The rab3 GDP/GTP exchange factor homolog AEX-3 has a dual function
 RT in synaptic transmission.";
 RL EMBO J. 19:4806-4816(2000).


```
CC -1- FUNCTION: GUANINE NUCLEOTIDE EXCHANGE FACTOR (GEF) FOR RAB3. MAY
CC REGULATE TWO DIFFERENT PATHWAYS FOR NEURAL ACTIVITIES.
CC -1- SUBUNIT: BINDS TO CAB-1.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF NEURONS.
CC -1- SIMILARITY: CONTAINS 1 DENN DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 093842; AAB52421.1; -.
CC DR EMBL: 049945; AAC47926.1; -.
CC DR WormRep: CO2H7.3; CE16806.
CC DR InterPro: IPR001194; DENN.
CC DR Pfam: PF02141; DENN; 1.
CC KW Guanine nucleotide releasing factor.
CC FT DOMAIN 216 364 DENN.
CC FT DOMAIN 604 632 SER-RICH.
CC SQ SEQUENCE 1409 AA; 157458 MW; 2DDE6395AC96313 CRC64;

Query Match 2.88; Score 106.5; DB 1; Length 1409;
Best Local Similarity 20.0%; Pred. No. 1.9;
Matches 115; Conservative 77; Mismatches 201; Indels 181; Gaps 25;

101 EAGKSESELOESNRASLOAOVYLSFOAPVIRACALLEVOVPADLYVQSGVSAVDFCE 160
178 QOSKNQOTVAFEDQALVQSGVSGMFSAFKLR--LMEDESLRELV-----CSK 825
161 ASLGAEOVQISYTKPRYOKELNLTQO-----LPDC-RGLEVRDSIQSC----- 202
826 LNLGLEVKL--SEDEVKEVQLTKGQKAYVKILAKLEGIEVSNTGCCGFASVHY 882
203 -----WVLPMLNVSTDDGNVLTLLDVSEODESFLLYLRVPDALKSLMYKNTGP- 253
883 LEIATHHYAMMGGEVTPSSASPTMTTPEHSNDILKESRPKLP-STDIDRTPTKPL 941
254 -ONITLNTD-----LVPCLCIOYWSLEPDSERVEFCPPREDP--GAHRLMHIARL 302
942 GQNVYTTSTNNHLEAOSTKSPALPPV---PPREAPPI-PKRNPPILGAPKVPBGAR- 995
303 RVLSPEWQDLADCCLPKGYTLQWAPQSP-----CQPLV 338
996 -----APPILPPRPKVKTTAVDEFPQNLVNNOPAPQSSPFLADADEQTRPL 1044
339 PPVPQKNATVNEQDQOLVAGHPR-----LCVQVSTWEKVOIQASWAD--- 382
1045 KRAPPTTLVQKQEPCK-VLPTPNEPVRYHYIQELILAVQHQIMONLQFWENAFVLDVAQ 1103
383 ---SLGPFKDDMLLVKKTGLNNTSVCALEPSGCTPLPSMASTRARL-----GEELDD 434
1104 EREIYGMADPEPMIDRYSLANDESEKRLLEEDRLLSTLLHNTMYMTMCGTGOKRALD- 1162
435 FSHQCMQMLMNDNGSLMACPMQKYIHRWVVLWACLILAALEFFLLKDRKKAAR 494
1163 ---QKVRLLGLKRAHIGLVKSEIKNLDE-----LPSTGQNFITPLK-----PL 1202
495 GSR-----TALLHSADGAG-----YERLVGALASALSO 523
1203 GSRLVQKQSEFTVCGQSSDQMMEMVEYCDVAVVLSITGAATERMYERLVN-----ITY 1257
524 MPLRVAVDLSRRELSAAGALWPHHQRRLIOE 557
1258 SPKTKILCLMRHDDKVH--MKKHTKKCKRELYO 1289
```

```
AC P23701;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 kDa outer membrane protein precursor (cysteine-rich outer membrane
DE protein) (CRP) (60 kDa cysteine-rich OMP).
GN OMCB OR OMP2 OR ENVB.
OS Chlamydia psittaci (Chlamydia psittaci).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EA6/A22/M;
RX MEDLINE=90384851; PubMed=2402464;
RA Watson M.W., Lambden P.R., Clarke I.N.;
RT "The nucleotide sequence of the 60 kDa cysteine rich outer membrane
RT protein of Chlamydia psittaci strain EA6/A22/M."
RL Nucleic Acids Res. 18:5300-5300(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=6BC;
RX MEDLINE=91267949; PubMed=2050637;
RA Everett K.D.E., Hatch T.P.;
RT "Sequence analysis and lipid modification of the cysteine-rich
RT envelope proteins of Chlamydia psittaci 6BC."
RL J. Bacteriol. 173:3821-3830(1991).
RN [3]
RP SEQUENCE FROM N.A.
RA Watson M.W.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES
CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL
CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT
CC VIRULENCE FACTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X53512; CA437592.1; -.
CC DR EMBL: M61116; AAB61619.1; -.
CC DR PIR: S12603; S12603.
CC DR InterPro: IPR003506; Chlam_OMP6.
CC KW PRINTS: PR01336; CHLAMIDIOM6.
CC DR Outer membrane; Transmembrane; Signal; Virulence.
CC FT SIGNAL 23 40
CC FT PROPEP 1 22
CC FT CHAIN 41 557 60 KDA OUTER MEMBRANE PROTEIN.
CC FT VARIANT 45 45 A -> S (IN STRAIN 6BC).
CC FT VARIANT 73 73 E -> G (IN STRAIN 6BC).
CC SQ SEQUENCE 557 AA; 59843 MW; 0D444F09EAA073C6 CRC64;

Query Match 2.88; Score 103; DB 1; Length 557;
Best Local Similarity 20.1%; Pred. No. 1.1;
Matches 61; Conservative 42; Mismatches 97; Indels 104; Gaps 15;

158 CF-EASLGAEOVQISYTKPRYOKELNLTQOLPCRGLEVDSIOSCVLTPVLIANTSDDN 216
198 CFTATVYACAPFELRSTYTKG-OPALCTKQDGPFCAR-----CPVYKIEVCNTGSA 249
217 VLLTLDVSEODESFLLYLRVPVDAKLSLYKNTLGPONTLNTDLPV-----CLCIQVW 272
250 IARNVVD-----NVPVDS-----YTHASGQVRLSFLNGDMRPGDSKQFC----- 289
273 SLEPDSERVEFCPPREDP-----GAHRLMHIARLVLSPGWQDLADP----- 316
290 -----VEFCPPORRGKVTNATVSYCGHGKCSANVTY-----VNEPCVQVNI 332
```


OY 317 -----CLPGKVTLCWQAP-DOSPCQPIVPPQKNAIVNEPOFQVAGHPNLCVOVS 368
 DB 333 GADMSVVCAPVETTVVSNPDGLKLVYVEDTAPSGATI-----LEAGAEICCNKA 385
 OY 369 TW-----EKVOAQCSWADSGAPFKDMLL-----VEKMT---GLNNTSV 405
 DB 386 VMCIKEMCGEITQIFVVAAGSPGKFTNOVVATNSDCGCTSCAEVYTHMKGLATM 445
 OY 406 CALE 409
 DB 446 CVID 449
 RESULT 7
 APAP_HUMAN STANDARD; PRT; 1248 AA.
 ID APAP_HUMAN
 AC 014727; Q9UBZ5; Q9UJ58; Q9UJ59; Q9UJ60; Q9UJ61; Q9UJ62; Q9UJ63;
 AC Q9UG64; Q9UG65; Q9UG66; Q9UG67; Q9UG68; Q9UG69; Q9UG70;
 AC Q9UG90; Q9UG91; Q9UG92; Q9UG93; Q9UG94; Q9UG95; Q9UG96; Q9UG97;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Apoptotic protease activating factor 1 (Apaf-1).
 GN APAF1 OR KIA0413.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OC NSL Taxid:9606;
 RN 11
 RP SEQUENCE FROM N.A. (ISOFORM 3), AND PARTIAL SEQUENCE.
 RC TISSUE=Cervical carcinoma;
 RX MEDLINE-97410306; PubMed-9267021;
 RA Zou H., Henzel W.J., Liu X., Lutschg A., Wang X.;
 RT "Apaf-1, a human protein homologous to C. elegans CED-4, participates
 RT in cytochrome c-dependent activation of caspase-3.";
 RT Cell 90:405-413(1997).
 RN 12
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 4 AND 5).
 RC TISSUE=Peripheral blood, Heart, and Cervical carcinoma;
 RX MEDLINE-99373149; PubMed-10441496;
 RA Hahn C., Hirsch B., Jahnke D., Duerkop H., Stein H.;
 RT "Three new types of Apaf-1 in mammalian cells";
 RT Biochem. Biophys. Res. Commun. 261:746-749(1999).
 RN 13
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=T-cell;
 RX MEDLINE-99292765; PubMed-10364241;
 RA Saleh A., Srinivasula S.M., Acharya S., Fishel R., Alnemri E.S.;
 RT "Cytochrome c and dATP-mediated oligomerization of Apaf-1 is a
 RT prerequisite for procaspase-9 activation.";
 RT J. Biol. Chem. 274:17941-17945(1999).
 RN 14
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Kidney;
 RA Benedict M.A., Nunez G.;
 RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN 15
 RP SEQUENCE OF 385-1248 FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain;
 RX MEDLINE-98116655; PubMed-9455477;
 RA Ishikawa K.-I., Nagase T., Nakajima D., Seki N., Ohira M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. VIII.
 RT 78 new cDNA clones from brain which code for large proteins in
 RT vitro.";
 RN DNA Res. 4:307-313(1997).
 RN 16
 RP SEQUENCE OF 810-864 AND 866-883 FROM N.A.
 RA Roberts D.L., Daigleish R., Cohen G.M., MacFarlane M.;
 RT "The mammalian CED4 homologue, APAF1, exists as two distinct forms in
 RT human cells.";
 RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN 17
 RP SEQUENCE OF 1-138 FROM N.A. (ISOFORM 1/4/5).
 RA Won M., Lee J.-W., Oh H.-H., Kim D.-U., Chung K.-S., Lee M.,
 RA Yoo H.-S.;
 RT "Cloning of variant Apaf1.";
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN 18
 RP APAF-1-MEDIATED OLIGOMERIZATION.
 RX MEDLINE-98315378; PubMed-9651578;
 RA Srinivasula S.M., Ahmad M., Fernandes-Alnemri T., Alnemri E.S.;
 RT "Autoactivation of procaspase-9 by Apaf-1-mediated oligomerization.";
 RT Mol. Cell 1:949-957(1998).
 RN 19
 RP INDUCTION BY E2F AND P53.
 RX MEDLINE-21283226; PubMed-11389439;
 RA Moroni M.C., Hickman E.S., Denchi E.L., Caprara G., Coll E.,
 RA Cecconi F., Mueller H., Helin K.;
 RT "Apaf-1 is a transcriptional target for E2F and p53.";
 RT Nat. Cell Biol. 3:552-558(2001).
 RN 20
 RP X-RAY CRYSTALLOGRAPHY (1.3 ANGSTROMS) OF 1-97.
 RX MEDLINE-20013059; PubMed-10543941;
 RA Vaughn D.E., Rodriguez J., Lazebnik Y., Joshua-Tor L.;
 RT "Crystal structure of Apaf-1 caspase recruitment domain: an alpha-
 RT helical Greek key fold for apoptotic signaling.";
 RT J. Mol. Biol. 293:439-447(1999).
 RN 21
 RP STRUCTURE BY NMR OF 1-97.
 RX MEDLINE-20047184; PubMed-10578182;
 RA Day C.L., Dupont C., Lackmann M., Vaux D.L., Hinds M.G.;
 RT "Solution structure and mutagenesis of the caspase recruitment domain
 RT (CARD) from Apaf-1.";
 RT Cell Death Differ. 6:1125-1132(1999).
 CC 1- FUNCTION: Oligomeric Apaf-1 mediates the cytochrome c-dependent
 CC autocatalytic activation of pro-caspase 9 (Apaf-3), leading to the
 CC activation of caspase-3 and apoptosis. This activation requires
 CC ATP.
 CC 2- SUBUNIT: Monomer. Oligomerizes upon binding of cytochrome c and
 CC dATP. Oligomeric Apaf-1 and pro-caspase-9 bind to each other via
 CC their respective NH2-terminal CARD domains and consecutively
 CC mature caspase-9 is released from the complex.
 CC 3- SUBCELLULAR LOCATION: Cytoplasmic.
 CC 4- ALTERNATIVE PRODUCTS: 5 isoforms; 1/Apaf-1XL (shown here), 2/Apaf-
 CC 1L, 3/Apaf-1S, 4/Apaf-1M and 5/Apaf-1XS; are produced by
 CC alternative splicing.
 CC 5- TISSUE SPECIFICITY: Ubiquitous. Highest levels of expression in
 CC adult spleen and peripheral blood leukocytes, and in fetal brain,
 CC kidney and lung. Isoform 1 is expressed in heart, kidney and
 CC liver.
 CC 6- INDUCTION: By E2F and p53 in apoptotic neurons.
 CC 7- SIMILARITY: CONTAINS 1 CARD DOMAIN.
 CC 8- SIMILARITY: CONTAINS 13 WD REPEATS (TRP-ASP DOMAINS).
 CC 9- SIMILARITY: CONTAINS 1 NB-ARC DOMAIN.
 CC 10- CAUTION: Ref.7 sequence differs from that shown due to a
 CC frameshift in position 109.
 CC 11- This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 DR EMBL: AF013263; AAC51678.1;
 DR EMBL: AJ243003; CAB55579.1;
 DR EMBL: AJ243004; CAB55580.1;
 DR EMBL: AJ243005; CAB55581.1;
 DR EMBL: AJ243006; CAB55582.1;
 DR EMBL: AJ243007; CAB55583.1;
 DR EMBL: AJ243008; CAB55584.1;
 DR EMBL: AJ243009; CAB55585.1;
 DR EMBL: AJ243010; CAB55586.1;


```

>R EMBL: AJ243011; CAB55587.1; -
>R EMBL: AJ243048; CAB55588.1; -
>R EMBL: AJ243107; CAB56462.1; -
>R EMBL: AF134397; AAD38344.1; -
>R EMBL: AF149794; AAD34016.1; -
>R EMBL: AB007873; BAA24843.1; -
>R EMBL: AJ133643; CAB55085.1; -
>R EMBL: AJ133644; CAB55086.1; -
>R EMBL: AF248734; CAB55087.1; -
>R PDB: 1CYS; 01-DEC-99.
>R PDB: 1CWM; 21-JAN-00.
>R MIM: 602233; -
>R InterPro: IPR001315; CARD.
>R InterPro: IPR002182; NB-ARC.
>R InterPro: IPR001680; WD40.
>R Pfam: PF00619; CARD; 1.
>R Pfam: PF00931; NB-ARC; 1.
>R Pfam: PF00400; WD40; 11.
>R PRINTS: PRO0320; GPROTEINBRPT.
>R SMART: SM00320; WD40; 10.
>R PROSITE: PS00678; WD_REPEATS_1; 4.
>R PROSITE: PS00678; WD_REPEATS_2; 9.
>R PROSITE: PS00294; WD_REPEATS_REGION; 1.
>R Apoptosis; ATP-binding; Repeat; WD repeat; Alternative splicing;
>R 3D-structure.
>R DOMAIN 1 90
>R NP_BIND 154 415
>R REPEAT 613 652
>R REPEAT 655 694
>R REPEAT 697 738
>R REPEAT 741 780
>R REPEAT 796 836
>R REPEAT 838 877
>R REPEAT 880 919
>R REPEAT 959 998
>R REPEAT 1001 1040
>R REPEAT 1042 1080
>R REPEAT 1083 1122
>R REPEAT 1125 1164
>R REPEAT 1175 1212
>R DOMAIN 95 98
>R VARSPPLIC 99 109
>R VARSPPLIC 575 575
>R VARSPPLIC 824 866
>R VARSPPLIC 1113 1154
>R CONFLICT 108 108
>R CONFLICT 134 134
>R CONFLICT 145 145
>R CONFLICT 161 161
>R CONFLICT 370 370
>R CONFLICT 383 383
>R CONFLICT 544 544
>R CONFLICT 580 580
>R CONFLICT 608 608
>R CONFLICT 620 620
>R CONFLICT 639 639
>R CONFLICT 708 708
>R CONFLICT 742 742
>R CONFLICT 746 746
>R CONFLICT 757 757
>R CONFLICT 795 795
>R CONFLICT 798 798
>R CONFLICT 825 825
>R CONFLICT 871 871
>R CONFLICT 876 876
>R CONFLICT 949 949

```

Query Match 2.7%; Score 100.5; DB 1; Length 1248;
Best Local Similarity 20.3%; Pred. No. 5;

Matches 118; Conservative 65; Mismatches 210; Indels 187; Gaps 29;

```

>Q 79 DCALRVAVVHLAVHGH-----VAPEEAGKSDSEIOE-SRNASIOA-QVYLSFOAY-- 128
>Db 537 DCVASENFQOEFLSLGHLGKQPPNIVQIGCEPSTSVYQDAKQAOEVNGLMYLE 596
>Q 129 -----PIARCALLLEVQVPADLV-----OPGQSVGSNVEDCEASIGAEQVMSYTPKRY 177
>Db 597 WINKKNTNLSRLVVRPHITDAVYHACFSEIDGRI-----ASGADKTLQVFAETG 647
>Q 178 QKELNLTQQLPD--CRGLEVRDS-IQSCVYLPALNV--STDGNNVLLTIDVSEEQ----- 227
>Db 648 EKLEIKAHDEEVLCGFSTDRFATCSVYKVKVIMNSMTGE-LVHTYDEHSEQVWCH 706
>Q 228 ---DFSLYLRVPYDALKSLMYKNLGPONTLHTDVLPCLOIOWSLPEPSENERFC 284
>Db 707 FTNSSHLLATGSSDCFLKMDLNQKRCRNTMFGHTN-----SVNHC 749
>Q 285 PFREDDPGAHRLMHLARLVLSPGVQOLDAPCLPGKVTYLCWQAPDQSPQVLPVPPQK 344
>Db 750 RFSPDD-----KLLASCSADGTLKL-WDATSANERKST--NVKQF 786
>Q 345 NAIYVNEPQDQVLVAGHRLVCYQVSTWEVYQAGSM-ADS-----LGPFRKDM 391
>Db 787 FLNLDEPQEDM-----EVIYVCCSWSDGARIWAAKNKIFLFDIHTSG 830
>Q 392 LLYEMKTGLNNT-SVCALPESGCTPLPSMASTRARLGEQLDQFRSHOCMLMNDNMG 450
>Db 831 LGEIHTGHSTIOYCPSPNNHLAVNL-----SOYCEVLEHNTDSRS 873
>Q 451 SLWACPDKTYIHRWVLTWLTACLLAALFFELLKDKRKARAGSRTALLHSADAGY 510
>Db 874 KVADCRG---HLISWV-----HG-----VHSPDGSF 897
>Q 511 ERLVGLASALSQPLVAVDVMRSRELISANGALAMFHQRRIIOEGVILLSPAAV 570
>Db 898 ---LTSSDDO---TRLMEKTKVCKNSAV-MLKQEVNVQVQENVMVL-----AV 940
>Q 571 AOCQOMLQIQTVEPGPHDAL--AAWLSGVLPDPLDGRATG 608
>Db 941 DHIIR-LQLINGRTGQIDYLTEAOVSCCCLSPHIOYIAGF 979

```

RESULT 8
A2M2_MOUSE
ID A2M2_MOUSE STANDARD; PRT: 1451 AA.
AC P28666;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, last sequence update)
DT 01-NOV-1997 (Rel. 35, last annotation update)
DE Murioglobulin 2 precursor (Mug2).
GN Mug2 OR Mug2-2.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
RN NCBI_Taxid=10090;
RX TISSUE= Liver.
RC MEDLINE=91358495; PubMed=1840592;
RA Overberg L., Torrekens S., van Leuven F., van den Berghe H.;
RT "Molecular characterization of the murglobulins.";
RL J. Biol. Chem. 266:16903-16910(1991).
-I- FUNCTION: A PROTEINASE ACTIVATES THE INHIBITOR BY SPECIFIC
PROTEOLYSIS IN THE BAIT REGION, WHICH, BY AN UNKNOWN MECHANISM
LEADS TO REACTION AT THE CYS-TEINYL-GLUTAMYL INTERNAL THIOL ESTER
SITE AND TO A CONFORMATIONAL CHANGE, WHEREBY THE PROTEINASE IS
TRAPPED AND/OR COVALENTLY BOUND TO THE INHIBITOR. WHILE IN THE
TETRAMERIC PROTEINASE INHIBITORS STERIC LINKAGE BETWEEN THE
STRONG, MONOMERIC FORMS NEED A COVALENT LINKAGE BETWEEN THE
ACTIVATED GLUTAMYL RESIDUE OF THE ORIGINAL THIOL ESTER AND A
TERMINAL AMINO GROUP OF A LYSINE OR ANOTHER NUCLEOPHILIC GROUP ON
THE PROTEINASE, FOR INHIBITION TO BE EFFECTIVE.

	Query Match	2.78;	Score 99.5;	DB 1;	Length 1451;
	Best Local Similarity	19.48;	Pred. No. 7.4;		
	Matches 77;	Conservative 47;	Mismatches 159;	Indels 113;	Gaps 17;
3Y	67 QTELVLNCPQ-----KTDCLARVVVVHLAV-----BGHMAEPEPA---102				
DB	454 ETSILPCNQHIVQAHFILKGDGLVELFYLYLMAQSSLIQGNTHQVEBPAPK 513				
3Y	103 GKSDSELQESRMASLQAVVLSFOATYIPARCALLQVQPADLVQPGSGVSAVFDFEAS 162				
DB	514 GNFDE-----IPVEFSMAPMAKMLITYILPDGEVI--ADSVNFEIEKCLRKK 559				
3Y	163 LGAEOVIMSTKPRYQKELNLTQOLPDCRGLEVDSDISQCVLPLVLNVTSDGDVLLITLD 222				
DB	560 VDLSTFS--SOSLPASQIRLQVTAASPQSLCGLRAYD--QS-----595				
3Y	223 VSEEDQSEFLLYLRVPVDALKSIMYKYLGPONTLNHTDLP-----CLCIGQW 272				
DB	596 -----VLLLKLP--EDELSPSWIYMLPMQO-----HNKPIPSSLSSEDRDCLITYSM 640				
3Y	273 SLEPDSERVCECPREDPGAHR-----NIMHIAIRLVLSPCVWQOLDACCLP--GVVTLCW 326				
DB	641 YAE---KHTDMVFHGRKEDVYRYRYEDMDLKAFTNLIKLPRIQCDSAPMSGPRGKFDLAF 697				

RESULT	9	NOTCH_BARE	STANDARD:	PRT: 2437 AA.
ID	NOTCH_BARE	STANDARD:	PRT: 2437 AA.	
AC	P46530.1			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Neurogenic locus notch homolog protein precursor.			
GN	NOTCH OR NOTCH1A.			
OS	Brachydanio reio (Zebrafish) (Zebra danio).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;			
OC	Cypriniformes; Cyprinidae; Danio.			
OX	NCBI_TaxID=7955;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Embryo:			
RX	MEDLINE:94128602; PubMed-8297791;			
RA	Bierkamp C., Campos-Ortega J.A.;			
RT	"A zebrafish homolog of the Drosophila neurogenic gene Notch and			
RT	its pattern of transcription during early embryogenesis.";			
RL	Mech. Dev. 43:87-100(1993)			
CC	-1- FUNCTION: IMPLICATED IN CELL FATE SPECIFICATIONS DURING			
CC	EMBRYO DEVELOPMENT. MAY BE INVOLVED IN THE FORMATION OF THE			
CC	NEURAL PLATE, NOTOCHORD AND BRAIN VESICLES.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED IN ALL CELLS IN PRECASTRULATION			
CC	STAGES. DURING GASTRULATION IS DIFFERENTIALLY EXPRESSED,			
CC	ACCUMULATING PREDOMINANTLY IN THE PRECHORDAL MESODERM AND			
CC	NOTOCHORD. AT THE END OF GASTRULATION, EXPRESSED ALONG THE			
CC	ANTERIOR-POSTERIOR AXIS INCLUDING THE DEVELOPING NEURAL PLATE			
CC	AND DIFFERENTIATING MESODERM. ALSO PRESENT IN THE DEVELOPING			
CC	BRAIN AND HEAD REGIONS.			
CC	-1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.			
CC	-1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.			
CC	-1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.			
CC	-1- SIMILARITY: CONTAINS 6 ANK REPEATS.			
CC	-----			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: X69088; CAA48831.1; -			
DR	HSP: P00740; IEDM.			
DR	ZFIN: ZDB-GENE-990415-173; notch1a.			
DR	InterPro: IPR002110; ANK.			
DR	InterPro: IPR000152; Asx_hydroxyl.			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR000742; EGF_2.			
DR	InterPro: IPR001881; EGF_Ca.			
DR	InterPro: IPR001438; EGF_II.			
DR	InterPro: IPR000800; Notch.			
DR	Pfam: PF00023; ank; 6.			
DR	Pfam: PF00008; EGF; 36.			
DR	Pfam: PF00066; notch; 3.			
DR	PRINTS: PR00010; BCBLOOD.			
DR	PRINTS: PR01452; NOTCH.			
DR	SMART: SM00248; ANK; 5.			
DR	SMART: SM00179; EGF_CA; 19.			
DR	SMART: SM00001; EGF_LIKE; 16.			


```

OY 430 ELIADFRSHOCQMLNDNDMGL--WACPMXKTHRRVWLWVLACLLAALFFPLLKKD 488
DB 2001 ELI---NCHADPNNAVDSGKSLHMAAAVNN-----VDAAV---VLTKNG 2039
OY 489 RRRKARSR--TALLSHADGAGYER---LYGALASA-----LSQMLRVAVDMSREL 538
DB 2040 ANNDLONKKEETPEFLAARECS--YETAKVLIDHNRDIADHDLDQIDPRDIHER----- 2092
OY 539 SANGALAMFHRRRIIOEGGVILLFFSPAAYAOQ--OOMLQLOVTPGPHDALAAMLSC 596
DB 2093 -----MHDDIYALTEEMVIVNRPPLSPPLCCPNYLGIRK--SPGNNTAKKTRK 2143
OY 597 VLDPFLOGRAGR-----YGV-----YFDGLHPDSVPS 626
DB 2144 PGKGVGSGKDSQKIRTRKKSGDKNGIMEGVLSPYDLSLSPHGYLSVSSSPMPT 2203
OY 627 PFRVAPFSLPTOLPAFLDALOGGCGSTSGRPRADREVRVT 666
DB 2204 PFGOSPIISL--NQLQGLADSHMGALQGLGKFPDSAPRUS 2242

RESULT 10
GLSK_HUMAN STANDARD: PRF: 669 AA.
AC 094925: 09UN40; 09UL05; 09UL06; 09UL07;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutaminase, kidney isoform, mitochondrial precursor (EC 3.5.1.2)
DE (GLS) (L-glutamine amidohydrolase) (K-glutaminase).
GN GLS OR KIAA0838.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RC TISSUE=Placenta;
RX MEDLINE=20473703; PubMed=11015561;
RA Elgadi K.M., Meguid R.A., Qian M., Souba W.W., Abcouwer S.F.;
RT Cloning and analysis of unique human glutaminase isoforms generated
RT by tissue-specific alternative splicing.
RL Physiol. Genomics 1:51-62(1999).
[2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=99156230; PubMed=10048485;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT Prediction of the coding sequences of unidentified human genes. XII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.
RL DNA Res. 5:355-364(1998).
[3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RA Chavez R.A., Wang C., Cong R., Hawkins J.E., Forsayeth J.R.;
RT Identification and expression of human renal and hepatic glutaminase
RT isoforms.
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=20185643; PubMed=10719215;
RA Holcomb T., Taylor L., Trochimoto J., Curthoys N.P.;
RT Isolation, characterization and expression of a human brain
RT mitochondrial glutaminase cDNA.
RL Brain Res. Mol. Brain Res. 76:56-63(2000).
[5]
RP SEQUENCE OF 225-466 FROM N.A.
RC TISSUE=Colorectal carcinoma;
RA Turner A., McGivan J.D.;

```

```

RT "Adenoma and carcinoma cell lines derived from colorectal tumours
RT express different isoforms of glutaminase."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE FIRST REACTION IN THE PRIMARY PATHWAY
CC FOR THE RENAL CATABOLISM OF GLUTAMINE.
CC -1- CATALYTIC ACTIVITY: L-glutamine + H(2)O = L-glutamate + NH(3).
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; KGA (SHOWN HERE), GAM AND
CC GAC; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: KGA IS EXPRESSED PREDOMINANTLY IN BRAIN AND
CC KIDNEY BUT NOT IN LIVER, GAC IS EXPRESSED PRINCIPALLY IN CARDIAC
CC MUSCLE AND PANCREAS BUT NOT IN LIVER OR BRAIN, GAM IS
CC EXPRESSED SOLELY IN CARDIAC AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE GLUTAMINASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 ANK REPEAT.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF158555; AB047056.1; -
DR EMBL: AF097492; AAF00088.1; -
DR EMBL: AF097493; AAF00089.1; -
DR EMBL: AF097495; AAF00090.1; -
DR EMBL: AB020645; BA074861.1; -
DR EMBL: AF223943; AAF33825.1; -
DR EMBL: AF327434; AAG47842.1; -
DR EMBL: AF279697; AAG17700.1; -
DR MIM: 138280; -
DR InterPro: IPR002110; ANK.
DR Pfam: PF00023; ank. 2.
DR SMART: SM00248; ANK_1.
DR PROSITE: PS50088; ANK_REPEAT; 1.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR KX Hydroxylase; Transit peptide; Mitochondrion; ANK repeat;
KW Alternative splicing
FT TRANSIT 1 16 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 17 669 GLUTAMINASE, KIDNEY ISOFORM.
FT REPEAT 585 618 ANK.
FT VARSPPLIC 162 169 ALKSTGLR -> VSFYILS (IN ISOFORM GAM).
FT VARSPPLIC 170 669 MISSING (IN ISOFORM GAM).
FT VARSPPLIC 551 669 VKSYINLFAVYTGVSALRPALSLAMMEQSDYDSRTALH
FT VAAEGHEVEVFKLEAKVNPFPDRMNITPDALHGH
FT HDVFKILQEVQVITPGDSDGKNGKQVTHKNDGL ->
FT HSFGLDYESLQDELALKEITWKVSPESNEDISTVYRM
FT ESTLEKRS (IN ISOFORM GAC).
FT G -> S (IN REF. 1; AAF00090).
FT E -> D (IN REF. 1; AAF00090).
FT F -> L (IN REF. 4).
FT V -> A (IN REF. 1).
SQ SEQUENCE 669 AA; 73461 MW; 4E5E63505E84E0B7 CRC64;

Query Match 2.68; Score 98.5; DB 1; Length 669;
Best Local Similarity 19.78; Pred. No. 3.2;
Matches 137; Conservative 67; Mismatches 219; Indels 273; Gaps 31;

OY 69 ELVLRCPQKTDCAI--RVVVVHLA-----VHGHW-----APEEAG 103
DB 12 DILLRSPAGVSATLRRQPLVTLGRRPGGRPAAGPAAARLHVMWGGGPAEPLARG 71
OY 104 KSISE---LOESRNASLOAOVLSFOAYPIARCALLEVOVPADLVOPGQSVGSAYVDCFE 160
DB 72 LSSSPSEILLQELGKSTHPQPGVSPAPAPAA-----PGKDGPGETDARG 116
OY 161 ASLGAEV-----QIMSYPKPRQKELNLNQQLPDCRGLFVRDS---IQ 200
DB 117 NSBCKELVASGENKIKQGLPLSLEDLFTYTAEGEKIPVHFITALKSTGLRTSDPRK 176
OY 201 SCWVLPMLNVSTGDGNVLTLLDVSEEDFSLYLRP-----VPDAIKSLMYKNLTGP 253

```


Query Match	Similarity	Score	DB 1	Length
Best Local	93	24.18	Pred. No. 12	1618
Matches	93	Conservative	45	Mismatches 139; Indels 109; Gaps 21
QY	349	NEPDEFQALVAGHNPILCYOV-----STWEK-----VOLQAC--SWADSLGPPKD---DM	391	
Db	83	NLAEELEGVAGR--CEQLRLARERTTEEVARNRRVAEAKCARAWLSOGAELEHELEA	139	
QY	392	LIV---EMKTGLMNTSYCA---LEPSGCTPLPSMASTRARALGEBELLQDFSHQCMQIW	444	
Db	140	LEVAHEEEVREGVLAQAACAPRLPAPRPPAPAPAEVEEL-ARRLEGAMRGAVRGYQ----	193	
QY	445	NDNNGSLWACPMDXIYHRRWVL----VWLCCLLALALFFLLILKDKRKAARG---SR	497	
Db	194	-----ERVAHMETSLDQTRERLRAAVQGAR---EVRLQLQLOLQERGLLER	237	
QY	498	TALLHSADAGYERL-----VGLASALSOMPLRVANDVLSRRELSAHGALAMFH	548	
Db	238	RALDORLEGRNORERLRATKFKFLQVALBALBEKQGLSQIAOVYLGKQOLA-----	288	
QY	549	HORRRILOGGVILLSPAAVAQCQOMQO-----LQTVBRPHALA---AWLSCLPLD	600	
Db	289	HLKMSLSLE-----VATYRFLLEANSRLQTPGGSGKSLSPQDDKLELQFPR	336	
QY	601	FLQGRATGKRYGVYFEGGLHPDVSPPFRVAAPFLSLPTQLPAFL---DALGGCSTSGR	657	
Db	337	TEEGRRIGSLTPV-----LSPTSLPSPLPA---TLTPVPAFLKNOEFLQARTPTLAST	387	
QY	658	PADRVERTQALRSALDSCSSSEAP	683	
Db	388	P---TPPTQAPSPAVDAETIRADAP	410	
RESULT	12			
APAF_RAT				
ID	APAF_RAT	STANDARD;	PRT;	1249 AA.
AC	O9EPV5;			
DT	01-MAR-2002 (Rel. 41, Created)			
DT	01-MAR-2002 (Rel. 41, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Apoptotic protease activating factor 1 (Apaf-1).			
GN	APAF1.			
OS	Rattus norvegicus (rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley;			
RA	Itch T., Itch A., Pleasure D.;			

EMBL: AF218840; AAF01013.1.
InterPro: IPR002934; NTP_transf.
InterPro: IPR001201; PAF_25A_core.
Pfam: PF01909; NTP_transf_2; 1.
mRNA processing: Transferase; Transcription: RNA-binding
Nuclear protein. 113
AC1_SITE 113 BY SIMILARITY.

RA "Cyclin T1 domains involved in complex formation with Tat and TAR RNA
GAYNOR R.B.: MEDLINE=99263518; PubMed=10329125;
RA Ivanov D., Kwak Y.T., Nee E., Guo J., Garcia-Martinez L.F.,
RA Gaynor R.B.: FUNCTION: 1

RT are critical for tat-activation.";
 RL J. Mol. Biol. 288:41-56(1999).
 RN [4]
 RP MUTAGENESIS.
 RX MEDLINE-99263519; PubMed-10329126;
 RA Kwak Y.T., Ivanov D., Guo J., Nee E., Gaynor R.B.;
 RT "Role of the human and murine cyclin T proteins in regulating HIV-1
 RT Tat-activation.";
 RL J. Mol. Biol. 288:57-69(1999).
 RN [5]
 RP FUNCTION.
 RX MEDLINE-99051315; PubMed-9832804;
 RA Garber M.E., Wei P., Kewalramani V.N., Mayall T.P., Herrmann C.H.,
 RA Rice A.P., Littman D.R., Jones K.A.;
 RT "The interaction between HIV-1 Tat and human cyclin T1 requires zinc
 RT and a critical cysteine residue that is not conserved in the murine
 RT Cyclin T protein.";
 RL Genes Dev. 12:3512-3527(1998).
 RN [6]
 RP FUNCTION.
 RX MEDLINE-99292873; PubMed-10364329;
 RA Bieleas P.D., Grdina T.A., Bogerd H.P., Cullen B.R.;
 RT "Analysis of the effect of natural sequence variation in Tat and in
 RT cyclin T on the formation and RNA binding properties of Tat-cyclin T
 RT complexes.";
 RL J. Virol. 73:5777-5786(1999).
 RN [7]
 RP FUNCTION: REGULATORY SUBUNIT OF THE CYCLIN-DEPENDENT KINASE PAIR
 CC (CDK9/CYCLIN T) COMPLEX, ALSO CALLED POSITIVE TRANSCRIPTION
 CC ELONGATION FACTOR B (P-TEF), WHICH IS PROPOSED TO FACILITATE THE
 CC TRANSITION FROM ABORTIVE TO PRODUCTION ELONGATION BY
 CC PHOSPHORYLATING THE CTD (CARBOXY-TERMINAL DOMAIN) OF THE LARGE
 CC SUBUNIT OF RNA POLYMERASE II (RNAP II). BINDS ALSO TO THE
 CC TRANSCRIPTIONAL DOMAIN OF THE HIV-1 NUCLEAR TRANSCRIPTIONAL
 CC ACTIVATOR, TAT. IT IS PROBABLY THE CIS-ACTING TRANSCRIPTION
 CC RESPONSE ELEMENT (TAR) RNA-BINDING COFACTOR FOR TAT. ALSO SERVES
 CC AS AN ESSENTIAL COFACTOR FOR HIV 2-TAT AND SIMIAN IMMUNODEFICIENCY
 CC VIRUS (AFRICAN MANDRILL) TAT FUNCTION.
 CC -1 SUBUNIT: ASSOCIATES WITH CDK9 TO FORM P-TEF. CYCLIN T1 IS THE
 CC PREDOMINANT CYCLIN ASSOCIATED WITH CDK9.
 CC -1 SUBCELLULAR LOCATION: Nuclear.
 CC -1 TISSUE SPECIFICITY: UNIKITOUSLY EXPRESSED.
 CC -1 MISCELLANEOUS: INTERACTION BETWEEN TAT AND CYCLIN T1 REQUIRES
 CC ZINC.
 CC -1 SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN C SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF045161; AAC39638.1;
 DR EMBL: AF048730; AAC39664.1;
 DR TRANSFAC: T02727;
 DR MIM: 602506;
 DR InterPro: IPR000553; Cyclin.
 DR Pfam: PF00134; cyclin; 1.
 DR SMART: SM00385; CYCLIN; 1.
 DR PROSITE: PS00292; CYCLINS; FALSE NEG.
 KW Cyclin: cell cycle; cell division; Colled coil;
 KW Transcription regulation; Nuclear protein.
 FT DOMAIN 253 270 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 254 272 TAT TAR RECOGNITION MOTIF (TRM).
 FT DOMAIN 384 425 COILED COIL (POTENTIAL).
 FT DOMAIN 517 526 POLY-HIS.
 FT DOMAIN 560 570 POLY-SER.
 FT DOMAIN 717 725 POLY-PRO.
 FT MUTAGEN 261 261 C->Y: LOSS OF HIV-1 TAT TRANSCRIPTION.
 FT CONFLICT 77 77 Q -> R (IN REF. 2).
 FT SEQUENCE 726 AA; 80684 MW; 4637EFB3DDEDEFJ3 CRC64;

Query Match 2.58; Score 93.5; DB 1; Length 726;
 Best Local Similarity 19.3%; Pred. No. 9.1;
 Matches 43; Conservative 39; Mismatches 70; Indels 71; Gaps 11;
 QY 50 PGSIQASAPGPLYVPRLOTELVLRCPORTDCALRVVYVLAHVGMAPEPEKGSDEL 109
 DB 79 PGN-SVAPALFLAKVEEQ-----PKLLEHYK--VAHTCLHPQESLPDT--RSEAYL 127
 QY 110 QESRNASLQAOVLSFOAYP-----IARCALLLEVOPADLVOPGSGAVFECFE 160
 DB 128 QOQVDVILESIITLTGLFELTIDPHRHVYKCTOL-VRASKDLAOTSTFMATNSLHLLT 186
 QY 161 ASLGAEOVWSYTRPRRQKEMLTQQLPDCRGLEVRDSIOSC-----VLLPMLN 209
 DB 187 FSL-----QYTPP-----VACVCHLACKSMWEIP--- 213
 QY 210 VSTDDG-----NVLLTIDVSEQDFSLILYLRPPVDALKSLM 246
 DB 214 VSTDGKHMEYVDATVTLLELDELTHEFLQILEKTPNRLKRW 256

Search completed: September 28, 2002, 19:43:38
 Job time: 286 sec


```

Db 267 GPAATFTLSEVETNEET-----PAG-----SLMRL-ETLLQSTQDSLPLPA 309
QY 442 -QLMNDMNGSLMACPMCKYIHRNV-----LWMLACILLAAALFFELLKKRRKAAGS 496
Db 310 EQAWNDD--GSL-----RRLMDRQPELLTELGRASRIFFELV-----PALRTA 351
QY 497 RTALLHSADGAGRYLVALGALSALSOMPLRYAVNDMSRRELSSAHGALAMFHQRRLIQ 556
Db 352 CPSLELDADA--YRFLSGTAVALDEAGEGVLLPSM-----W--DRRRL- 393
QY 557 EGVVILLFSP--AAVACOGMLQSTVE-----PGPHDALAAMLSCVLP----- 599
Db 394 -GLVLSAYTVDGVYGGASKRGKQOLVDFRMLAVGDDPLSEETALLETKSPILRLR 451
QY 600 -----DFOGRATGRYGVYFPDGL--HPSVSPFRRVAPLFSLPQLPA 642
Db 452 GOWVALDTEQMRKGLFETKPTGRKTAETLALASHHPDDVDTPLEVTAV-----RADG 506
QY 643 FLDALGGCSTAGRPADRVERTQALRSALDSCTS-----SEAPGCCCEEMDGPCTTL 697
Db 507 WLDGLLGAALASLQPLDPPDGFATLRPYQQRGLAMLAFLSLGLGSLADMDGLKTV 566
QY 698 E 698
Db 567 Q 567

RESULT 2
A54100
Tumor suppressor protein DCC precursor - human
N:Alternate names: colorectal cancer suppressor DCC
C:Species: Homo sapiens (man)
C>Date: 02-Aug-1994 #sequence-revision 02-Aug-1994 #text-change 05-Nov-1999
C:Accession: A54100; A40098
R:Hedrick, L.; Cho, K.R.; Fearon, E.R.; Wu, T.C.; Kinzler, K.W.; Vogelstein, B.
Genes Dev. 8, 1174-1183, 1994
A:Title: The DCC gene product in cellular differentiation and colorectal tumorigenesis.
A:Reference number: A54100; MUID:95011532
A:Accession: A54100
A:Molecule type: mRNA
A:Residues: 1-1447 <HED>
R:Fearon, E.R.; Cho, K.R.; Nigro, J.M.; Kern, S.E.; Simons, J.W.; Ruppert, J.M.; Hamilton
Science 247, 49-56, 1990
A:Title: Identification of a chromosome 18q gene that is altered in colorectal cancers.
A:Reference number: A40098; MUID:90100559
A:Accession: A40098
A:Molecule type: mRNA
A:Residues: 1-750 <PRA>
A:Cross-references: GB:M32292; NID:g181492; PIDN:AAA35751.1; PID:g181493
C:Genetics:
A:Gene: GDB:DCG
A:Cross-references: GDB:119838; OMIM:120470
A:Map position: 18q21.1-18q21.1
C:Keywords: transmembrane protein; tumor suppressor
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-1447/Product: tumor suppressor protein DCC #status predicted <MAT>

Query Match 3.0%; Score 112; DB 2; Length 1447;
Best Local Similarity 18.8%; Pred. No. 2.2; Indels 302; Gaps 34;
Matches 136; Conservative 74; Mismatches 210;

QY 20 STERLMEPODARCSLSCHLMDGVLCPLPSLQSPAG-PVLVPTRLQTEVL-----RC 74
Db 39 ALNFLSPSPDAVT-----MRGNNV-LDCSASDRGVPVTKKKDGIHIALGMDERK 89
QY 75 PORTDCALRVVVVLAHVGHMAEPE-----EACKSDSELOESRNASL-----Q 118
Db 90 QQLSNGSLILIONILHSRIH-----KPDGLYOCESLSDSGSIISRAKAVAVAPLRLSLQ 145
QY 119 AQVYLSQAIPI-ARCALLLEVQVA-----DL-----VQPG 148

```

```

Db 146 TESVTAFMGDTVLKCEVTEGEMPTIHMOKNQODLTPPIPGDSRVVVLPSGALQISRLQPG 205
QY 149 QSYSAVAFDC-----FEASLGAEOIWSYTKRKYKELNLTOQLPDCRGLFVNDISQSCW 203
Db 206 -DIG--IYKCSARNPASPRTGNEAEVRLSDPGLRHQLTFLQRPNSVVAIEGDAVLECC 262
QY 204 V-----LPLANVSTDDNVLTLIDVSEEDQFSEFLYLRPVDAKLSMYKNLTGPONI 256
Db 263 VSGPSPSPFWLR-----GEVY-----QKRSKYSILGGSNLL 236
QY 257 TLNHTDLPCLCICIOVWSLEPDSERVEFPFRRDPAN-----RNLMIHARLVLSPG 308
Db 297 ISNVT-----DQSMYTCVVYTKMENISASAEVLVLP- 330
QY 309 VMOLDAPCLPGKVTLCWQAPD-OSPCQPLVPPV-----QKNATVNEPODFOLVAGHPNL 363
Db 331 PWTLNHPSNL-----YAESMDIEFCTVSGKPVPTVMKMGDVYTPSDYQIVGGSNL 385
QY 364 CVQVSTWEKVQLQACSWADSLGPFKRD-----MLVEMKTLNNTSVCALEPSCGP- 415
Db 386 RI-----LGVAKSDGEFYQCVAVENEGAGNQTSAQLIVPKRAIPSSSV 427
QY 416 -----LPSMASTRARLGEELLQDFRSHQCMQMDNDNGSLMACPMCKYIHRNV 467
Db 428 LPSAPRDVPEVLVSSRRVLS-----WRPPE- 454
QY 468 VWLACILLAAALFFELLKKRRKAAGS-RTALLHSADGAGERYLVGALSALSMPL 526
Db 455 -----AKNINQTFYTFPSREGNRRERALTTPQGSQQLV 489
QY 527 R-----VAVDMSRRELSSAHGALAMFHQRRLILOEGGVILLFSPAAVACQ 574
Db 490 GNLKPEAMTYFRVAVAYNDMGESS-----QPIKVA-TQ 522
QY 575 QWQ-----LQVEEGPHALAMLSCLVPDFLQGR-----ATGRVGVYRD 616
Db 523 PELQVQEPVENLQAVSPTSILITWEP--PAYANGPVQGRFLCTEYSTGKEONIEVD 579
QY 617 GL 618
Db 580 GL 581

RESULT 3
T42214
Otogelin - mouse
N:Alternate names: mucin-like extracellular matrix protein
C:Species: Mus musculus (house mouse)
C>Date: 03-Dec-1999 #sequence-revision 03-Dec-1999 #text-change 17-Mar-2000
C:Accession: T42214
R:Cohen-Salmon, M.; El-Amraoui, A.; Leibovici, M.; Pett, C.
Proc. Natl. Acad. Sci. U.S.A. 94, 14450-14455, 1997
A:Title: Otogelin: A glycoprotein specific to the acellular membranes of the inner
A:Reference number: Z22079; MUID:98070772
A:Accession: T42214
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2910 <COH>
A:Cross-references: EMBL:U96411; NID:g2760883; PID:g2760884; PIDN:AB96561.1
A:Experimental source: strain BALB/c
A>Note: component of all the acellular membranes of the inner ear
C:Superfamily: von Willebrand factor; von Willebrand factor type A repeat homology;

Query Match 3.0%; Score 111.5; DB 2; Length 2910;
Best Local Similarity 19.8%; Pred. No. 6;
Matches 83; Conservative 43; Mismatches 140; Indels 153; Gaps 19;

QY 139 QVPADLVQGGQ-----VGSVAFDCFEASLGAE--VQIM-STTKRY 177
Db 2267 OVSSLSLSEQTRFRPDSACATADCSPLRLMVSNTFTSACHSFVSPESFCELMIRDTKYQ 2326

```


QY 178 OKELNLTQOLPPCRGLEV-----RDSIQSCWVLPWLVNSTDGDNVLLTLD 222
 Db 2327 OPCVALITYVAMCHKHCVCLEMGGSDYCPFLCSDSDTYQACVAACEPPTCODGVLPID 2386
 QY 223 VSEODE-----SFLLYL-----PYDALKSLMYKNLTPQNT 257
 Db 2387 PQCCVILGEGCVCTEETILHRRHSALCIPEDKCACTDSTGPRALGETNSSLG----- 2441
 QY 258 LNHDTLVPLCLTQVMSLEPSS-----ERYEFCFPREDPGAHRLMHTARLVLSR 307
 Db 2442 -----CCOQOQCAEDPTILRBLDCCPFPRESCP-----RFGVILLDP 2479
 QY 308 GWMQDAPCCLPCKVYLTCMAPOSPCQPLVPYPPQKNATVNEPQDFQLVAGH-----PN 362
 Db 2480 -----TEDPCLL-GSVVC-----NFTLCEGLAPTCRPGHSLTHFEQEDSCPSYCECDPG 2530
 QY 363 LCVQSTWKEVOLQACSMWDSLGPFRDMLLVEKMTGLNNTSVCLPESGC-----TPLS 418
 Db 2531 LC-----EAEVPTC-----REDQILLIEGRIG-----DSCCTSYRCGCCCECDPMPE 2572
 QY 419 MASTRA-----ARIGEELADFRSHQ-----CMQWMDNDMGLMACPMD 458
 Db 2573 COGEALVHHRMTTELCCPLVCCEHFRCPOVCGMGTSLVEVWSPRCPCPKSCCED 2631

RESULT 4

JEO347
 hypothetical protein c31612.1 (K1AA0734 (C2 domain protein)) [imported] - human
 N:Alternate names: BAI-associated protein 3
 C:Species: Homo sapiens (man)
 C>Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 01-Dec-2000
 C:Accession: JEO347; T45060
 R:Shirtsuchi, T.; Oda, K.; Nishimori, H.; Suzuki, M.; Takahashi, E.; Tokino, T.; Nakam
 Biochem. Biophys. Res. Commun. 251, 158-165, 1998
 A:Title: Cloning and characterization of BAP3 (BAI-associated protein 3), a C2 domain-co
 A:Reference number: JEO347; MUID:99009239
 A:Accession: JEO347
 A:Molecule type: mRNA
 A:Residues: 1-1187 <SHI>
 A:Cross-references: GB:AE017111; NID:g3929916; PIDN:BAA34710.1; PID:g3929917
 A:Experimental source: brain
 R:Frankland, J.
 submitted to the EMBL Data Library, July 1999
 A:Reference number: Z22901
 A:Accession: T45060
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 36-581, 'D', 583-1187 <FRA>
 A:Cross-references: EMBL:AL031709; PIDN:CAB56182.1
 A:Experimental source: clone LA16-316G12
 C:Genetics:
 A:Gene: BAP3
 A:Map position: 16p13.3-16p13.3
 A:Intons: 79/2; 108/3; 135/3; 171/3; 191/3; 214/1; 279/2; 294/2; 320/2; 354/3; 390/3;
 976/3; 1003/3; 1044/1; 1085/2; 1154/1
 A:Note: c31612.1

Query Match 2.9%; Score 107.5; DB 2; Length 1187;
 Best Local Similarity 21.0%; Pred. No. 4;
 Matches 158; Conservative 85; Mismatches 216; Indels 295; Gaps 47;

QY 1 MPVS-----WFLI-----SLALGRNPVVSLEKIMEPQDARCSIG-----LSCHL 41
 Db 356 VPVAGVDHMFLEPRSSASRVGCHLVLT--ITTORDTAMSGRGSGFLSHLLLSHL 413
 QY 42 -----WDGD-----VLCLPGSIQASAPGVLPVRLQTELYLKRCQKQKDC 80
 Db 414 LRLHSAEPNSSSNGELSTPAATILCLGA--QSNLSPLQLAV-LHWQVSSRHQ--TC 469
 QY 81 ALRVVVVHLA--VHGHNAPPEAGSKSDSELQSRNASLQA-----QVVSFOAY-PLA 131
 Db 470 TLDSYLLGLEDMQAHN--EAPSLPQOEESLADSLASFSEFGQLLKLQRLDYFPAT 526

QY 132 RCA-----LLEQVPADLVDPQGSVSAVFC-PEASLGAEV-----QIW 170
 Db 527 NSTAVHRLTLKCLCKLQLPQSPFEI-----CPFESELMNDIAALKRGKREMYDIL 580
 QY 171 SYTKRYQKELNLTQOLPPCRGL-----EVRDSIQSC----- 202
 Db 581 NAKSPREQPG-----PQRLP-----GLVVLADAVYDDLOFCYSVYASLFSHSLNVDVFLTRQ 634
 QY 203 -----WVLPMNVSTDGDNVLLTLDYSEEDPFFLLYLRVPDALKSLMYKNLTP 253
 Db 635 LERLVAEAMVL-----TEELSPKMTLEVASG--LFEYLTL-PLAD-LQRFW-DSICR 682
 QY 254 ONITLN-----HFDVPLCLQVW-----STEPDERVEFCFPREDP 290
 Db 683 DSRSLIAGINHAPFLP--AVKLMFOYLRODQAKRRLQGVADMDTLEP-----VDASSRSS 736
 QY 291 GAHNRIM--HIALRVLSPGWQDAPCCLPGRVTLQWAPDQSPCQPLVPYPPQK--NA 346
 Db 737 AATAGLCISHIOEL-----W-----VRLAM--PDPAOAGLQGLQGDVCEA 776
 QY 347 T-----VNEPQDFQVAG--HPNLCVQVSTWEKV-----LQACSMWDSLGPFRD 390
 Db 777 TLFYTELLKKKVDQVTOGGAAGEAVSEALCVLVNVEIVRKAAQALKGLAMDEG----- 829
 QY 391 MLVEMKTGLNNTSVCLPESGCTPLPSMASTRA-----ARIGEELADFRS 437
 Db 830 -----ATPEGVLPRLPLSCTQALDDLDQREAITVYHLSKAVGDIRK 873
 QY 438 H-QCMQWMDNDMGLMACPMDKYIHRRWLVMLACLLAALFFFLKKRRKARAS 496
 Db 874 YVOHISLSPDSIQNDEAVAPIMKYLDEK-----LALLNANSL-----VGN 913
 QY 497 RTALLHNSADGGERVELVAGALSALQSOMPLRAVADLMSRELSHGALAMHHORRRILQ 556
 Db 914 LSRVL-----EALWEILLQATILLQALG-ANRQVADYFSRPHFTLEALVSEPH-----A 960
 QY 557 EGVVILLFSPAAVACQOQWMLQOTVEPQPHDL 590
 Db 961 EG-----QGLPLESIRDSQYKRL 978

RESULT 5

JEO347
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: D75564
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, F
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1
 A:Reference number: A75250; MUID:20036896
 A:Accession: D75564
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-478 <WHI>
 A:Cross-references: GB:AE001869; GB:AE000513; NID:g6457721; PIDN:AAF09658.1; PID:g6
 A:Experimental source: Strain R1
 C:Genetics:
 A:Gene: DR0061
 A:Map position: 1

Query Match 2.8%; Score 106.5; DB 2; Length 478;
 Best Local Similarity 23.2%; Pred. No. 1.5;
 Matches 118; Conservative 38; Mismatches 184; Indels 169; Gaps 23;

QY 237 PVP-----DALKSLMYKNLTPGQNTLNHTDLYVCLCTQVMSLEPDSERVEFCFPREDGAH 293
 Db 17 PVPAGDVLVSLGNSVVRQIARTLGLLLLPAGAAQVQWA--LDRVEDDPPPPQGGD 73


```

OY 294 RNLHIAIRL-----VLSGV-----310
DB 74 ALFAOIAITROGSGSPHERAOLEVILRAGAPQSLRADPRERPLIREAORIRMLARR 133
OY 311 QLDAPCCLEPKVTLCWAPDQSPQPLVPVPOKNAVNEPODFOLVAGHPNLCOVSTW 370
DB 134 LARAPQOLPCTVAHSVLYAOEASALSAPAGDGLTLH-----LGSPD---RLAHV 182
OY 371 EKVOLACGMSADSLPCKMILVEMKTG---LNTSVCALESQCTPLPFSMASTRARLG 428
DB 183 OGRL---SR-SLGP---NMOLVQDETSDPVSQGLALTRPD---LP---ASER-----225
OY 429 BEILODFRSHOCQOLWMDNNGSLMACPMCKYIHRHRYLVWLACLLIAALFFELLKXD 488
DB 226 -QLFLVGGQOLQVLSGD-----YVLLRR 250
OY 489 -----RRKARSGSRVALLHSADAGYERLVGALASALQMPLRVAVDLMSRELS 539
DB 251 ADAAOTRHILARIALGRACALLLPAGOHGRILARILARRGDP-----PRADDP 302
OY 540 AHGALAMFH-----HORRILQBGVYVILFSPAANAQCOQWLOLQVPEPBDALAAW 593
DB 303 ASGOVAFAQTPAECLAAARSLORLEPLGRFSPAQVA-----HEAQVAQ 347
OY 594 LSCVLPDFQGRATGRVGVYFDGLHPDSVSPSPFRA---PLFSL---PQLPA---FL 644
DB 348 ALHLHP-----HYAAOVQAAEHAASNAEFLREANPDLVGREAGPVLPATGHFI 399
OY 645 DALGGCSTAGRPADPVERVTOALRSAL 673
DB 400 VFPLGPDPLEVRPLGDRLLTLRPDYRAEL 428

RESULT 6
T37188
presynaptic activity regulator aex-3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Feb-2000
C:Accession: T37188
R:Leimmac, D.; Minx, M.
Submitted to the EMBL Data Library, February 1996
A:Description: The sequence of C. elegans cosmid CO2H7.
A:Reference number: Z20523
A:Accession: T37188
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1409 <LEI>
A:Cross-references: EMBL:U49945; PIDN:ANC47926.1; GSPDB:GN00029
A:Experimental source: strain Bristol N2; clone CO2H7
C:Genetics:
A:Gene: aex-3; CESP:CO2H7.3
A:Map position: X
A:Introns: 77/1; 136/2; 183/2; 232/3; 283/3; 386/3; 427/3; 527/3; 577/1; 699/3; 782/2; 8

```

```

Query Match          2.8%; Score 106.5; DB 2; Length 1409;
Best Local Similarity 20.0%; Pred. No. 6;
Matches 115; Conservative 77; Mismatches 201; Indels 181; Gaps 25;

```

```

OY 101 EAGSDSELQSRNASIAQAVYLFQAYPLARCALLQVYPADLVQGGQSVGSAVFDFPE 160
DB 780 QOSNQGVKRCDDALVQSGVGMSPAKLR--LMEDSLRELIV-----CSK 825
OY 161 ASLGAEOIWSYTKPRYOKELNLTQO-----LPDC-RGLEVRDSIOSC-----202
DB 826 LNLGLEVL-----SEDEYVVEVLTQGFAYVYKILKACLEGIEVSFTPGCCGASVYHV 882
OY 203 -----WVLPMLVSTDGDNVLLTLDVSEODESFLYLRLVPDALKSLMYKNLTGP- 253
DB 883 LEIAHTHYWAGGGGFEVITPSSSAPSTMTTPSEHSNDILKESRKLPA-STIDLRTPTKPL 941
OY 254 -QNTILNHTD-----LVPLCLOVMSLEPDSEREPDFREDP--GAHRNIMHIARL 302
DB 254 -----LVPLCLOVMSLEPDSEREPDFREDP--GAHRNIMHIARL 302

```

```

DB 942 GQNTPTSTNNHIAOSTRSPALPPV-----PPREAPPI-PRKNPPGLGAPKEVGAR- 995
OY 303 RVLSPGVWOLDAPCCLEPKVTLCWAPDQSP-----COPLV 338
DB 996 -----APPLPRKRVKVTAVDERPONLVPNQNPQAPSSPSFLADADEQTPLL 1044
OY 339 PVPQKNAVNEPODFOLVAGHPN-----LCVOVSTWEKVOLOACSMAD-----382
DB 1045 KPAPPTTLFVGKQEPCK-VLPTNEPVRHYTYQGLILAVQIQWQNOQJEWENAFVDLVNQ 1103
OY 383 ---SIGPFRKDDMLVEMKGIANTSVCALEPSQCTPLPFSMASTRARL-----GELLAD 434
DB 1104 EREIVGMDPESEMIDRYSLANDSEKKRLLEEDRLSTLLHNTAVYIMCGTGOKALO- 1162
OY 435 FRSHOCQOLWMDNNGSLMACPMCKYIHRHRYLVWLACLLIAALFFELLKDKRKAAR 494
DB 1163 ---QVRRILGRKAHIGLVCSKEINKLDE-----LPSTQGNFPLK-----PL 1202
OY 495 GSR-----FALLHSADGAG-----YERLVGALASALSQ 523
DB 1203 GSRIVQKOSFTVCPGQSSDQGMFMFVCDVAVYLRSTIGATERMWERLVN-----ITY 1257
OY 524 MPLRYAVDMSRRELSAHGALAMFHQRRILQE 557
DB 1258 SPKTKIILCMRRHDKVH--MKRFHTKKRELQY 1289

RESULT 7
A38442
Probable tumor suppressor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 18-Oct-1991 #sequence_revision 18-Oct-1991 #text_change 16-Feb-1997
C:Accession: A38442
R:Nigro, J.M.; Cho, K.R.; Fearon, E.R.; Kern, S.E.; Ruppert, J.M.; Oliner, J.D.; Ki
Cell 64, 607-613, 1991.
A:Title: Scrambled exons.
A:Reference number: A38442; MUID:91121517
A:Accession: A38442
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-267 <NIG>
A:Cross-references: GB:M63696; GB:M63700; GB:M63702; GB:M63718; GB:M63698
C:Keywords: transmembrane protein

```

```

Query Match          2.8%; Score 106; DB 2; Length 267;
Best Local Similarity 24.3%; Pred. No. 0.77;
Matches 73; Conservative 35; Mismatches 100; Indels 92; Gaps 17;

```

```

OY 155 VFDCFEASLGAEOIWSYTKPRYOKELNLTQOLPDCRGLEVRDSIOSCWLPMLNVTSDG 214
DB 8 LYQC-EASLQSGSIIISRT-----AKAVAPLRFSLQTESVIAFM-----G 47
OY 215 DNVLLTLDVSEODSFLLYLRLVPDALKSLMYKNLTGPONTILNHTD-----LVPLCLOQ 270
DB 48 DTVLLKCEVIE-----PMPTI---HMOKN---QODLPIPDSDRVVVLPSGALQ 91
OY 271 VMSLEPDSEREPDFREDPGHRNIMHIARLVLSPGWQOLDAPCCLEPKVTLCWAPD 330
DB 92 ISRLQPDIDIGIRCSAR-NPASSRT-GNEAEVRIISP-PWFLNHPSTNL-----YAYESMD 143
OY 331 -QSPCOPLVPPVP-----QKNATVNEPODFOLVAGHPNLQVYSTWENVQLOACSMADSLG 385
DB 144 IEFECTVSGKVPVYVNMKNGDVYIPSDFIQVGGSNLRI-----LG 185
OY 386 PFKDD-----MLVEMKGIANTSVCALEPSQCTPLPS-----MASTRARL 427
DB 186 VKSDEGFYQCVAENENAGNAQTSAGLIYVK---PIPSSVLPSPADPVDVYVVLVSSREVR 242

RESULT 8
B39439
60k cysteine-rich outer membrane protein 1 precursor - Chlamydomonas reinhardtii

```


Db 1459 ELFERGTEVAGSLRVLRLDVAADDPERTIGSVLLDATERHRLVGLN--TTDG----- 1511

QY 221 LDVSEEDSFLL--YLKRPVDPAL-----KSLMYKNL-----TGPNIT 257

Db 1512 -PLPERETLTLFEQQAATPDATLVMGDRSLTVAELDARANRLAHLVALGAGPEQIV 1570

QY 258 L---NHTDLPVLCIOYWS-----LEPDSERVECFREDPGAHRLNHLIARLRLS 306

Db 1571 ALQDPRSLDLVTAN-LANWAGCAALYLPDPD-----YPPER-----IAHMRADA 1613

QY 307 PGWQDLAPCLPCKGVTLQWCA--SPDQSPCOPLVPVP---OKNATVNEPODFOLVA 358

Db 1614 RPLALVVA---IPDDTTLMAVADSRLTDADRSAPLLPAHPAVITYTSGSGAPRG--VVV 1668

QY 359 GHPLVCA-----QVSTWERYQLQACSW-----ADSLGPF 387

Db 1669 AHSILATVPAAQAAGLCTHSRLVNFASISPDAAVMELTSLAFTGAGLVADADLLPG 1728

QY 388 KDDMLVEMKGTGLNNTSVCALEPSCGCTPLPSMA 420

Db 1729 PSIALRVHNR---HTLIALPSPALPAPDGA 1757

RESULT 13

A82593

hypothetical protein XF2169 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: A82593

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequences

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MIMD:20365717

A>Note: for a complete list of authors see reference number A59328 below

A:Accession: A82593

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-698 <SIN>

A:Cross-references: GB:AE004030; GB:AE003849; NID:99107292; PIDN:AAF84968.1; GSPDB:GN001

R:Experimental source: strain 9a5c

R:Simpon, A.J.G.; Relnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Brlones, M.R.P.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, F.; Neto, E.; Docena, C.; El-Dorry, H.; Facinani, A.P.; Ferreira, A.J.S.

A:Submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Perro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitzima, J.P.; Krieger, J.E.; Kurama, E.E.; Laigt chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, F.A.; Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF2169

Query Match 2.7%; Score 101.5; DB 2; Length 698;

Best Local Similarity 19.3%; Pred. No. 6.1;

Matches 109; Conservative 66; Mismatches 197; Indels 193; Gaps 27;

QY 176 RQKELNTQQLPDCRGLEYRDSIQSCW-----VLPWLVNSTDGDNVLTLDV-- 223

Db 249 RLTFGEPELLAGHPD-RALTHLRQVLAWPEDPRTLEALLTAMKQDMD--DARITIDAL 306

QY 224 -----SEEDSFLLYLKRPVDPALKSLMYKNLGPONITLHTDL 263

Db 307 DLKPRHDLMLARLAAPVSGDEARIVIERLTSAMPEHLPAL--ETLMSLHDIQ--NNPEA 363

QY 264 VPCLCIOVWSLEP-----DSERVECFREDPGAHRLNHLIARLRLS 319

Db 364 AETVARQIVAIERGRISGEGRIVAEALLQDRPPT-----AAVACL- 401

QY 320 GKVTLCWQAPDQSCQPLVPVPQKNAVNEP-----QD-----FOLVAGHPNLQVSTW 370

Db 402 -----QOLIESVPEHERITLRWMLGLVQDRAGQFE-----ALATW 437

QY 371 EKVOLOACSWADSGFPFDDMLVEMKGTGLNNTSVCALEPSCGCTPLPSMASTRARLGE 430

Db 438 LOFOAEQAKYRLPLP-----QFTWNSKO-----WPDIAIPAEILARP 476

QY 431 LLQDFRSHQCQLNNDNMGSLMACPMDKYIHRMVLVYLACLLAALFFLLKKDR 490

Db 477 LL-----TWPGSHVERVYTNMAVASSMLC----- 502

QY 491 KAARGSTALLHNSADGAYRGLVAGALASALSOPLRVAVDLSRRELASNGA----- 543

Db 503 -ADRYSETP--PADPLQRYETVELS--GLTRPQALIDAM--RQLPARGIEDGNVID 554

QY 544 -LAFHHQ-----RRILQEGGVILLFSPAAVACQOQLOTOVEGPDAL--AAVLS 595

Db 555 WLLMWDNSLLTALPRLPREGRLITVLRPROM--LLDWIAYGSPFLADSLQQAANWLG 612

QY 596 CVLPDFLOGRATGRVGYFFDGLHPDSV--PSFRVA-----PLFSLPQLPAFLALQGG 650

Db 613 -----DILNQIALHLEDLPHHLIRDCIEDNPQALATTLIEDIFGSPFPPLPSLEAPR-- 666

QY 651 CSTSAGRPADRVREVTQALRSALDS 675

Db 667 -----LPAGRWRDYREVLSAFDA 685

RESULT 14

T04018

hypothetical protein F17A8.60 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999

C:Accession: T04018

R:Byvan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, I.

A:Submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15184

A:Accession: T04018

A:Molecule type: DNA

A:Residues: 1-1274 <BEV>

A:Cross-references: EMBL:AL049482

A:Experimental source: cultivar Columbia; BAC clone F17A8

C:Genetics:

A:Map position: 4

A:Introns: 55/3; 801/1

A>Note: F17A8.60

Query Match 2.7%; Score 101.5; DB 2; Length 1274;

Best Local Similarity 20.2%; Pred. No. 13;

Matches 112; Conservative 62; Mismatches 183; Indels 197; Gaps 33;

QY 187 LPDCRGLEYRDSIQS-----CWVLPWLVNSTDGDNVLTLDVSEEDSFLLYLRL 236

Db 694 LPEHFGRRKRDISSIVDRIRQSHSWSIRFL--SSAKQILKLAVALSSMSYAAACFKL 751

QY 237 PV-----PDALKSLW--YKNLTGPQNT--TLNHTDLVPCLCIOVWSL- 274

Db 752 PSLCKOIOSVLTFRFWWDSPKDRKMAVSWDKLTLPINEGGLGPRETEAKLS--WRL 808

QY 275 -EPDS--ERV-----EFCFREDPGAHRLNHL--IARLVLSG----- 308

Db 809 KPHSLLSVLLAKYNTSSFFMDCASPSASHGWCIGILGRDLIRKLGISIGOGDSIN 868

QY 309 WQDLAPCLLPKGVTLQWCA--SPDQSPCOPLVPVP--PKNATVNEPODFOLVAGHPNLQVOY 367

Db 869 VW-----TEAWLSF--SSPOTPIGPPTTNDLSV-----HDLICHDV 904

QY 368 STM--EKVOLOACSWADSLG-----PFRDMLVEMKGTGLNNTSV--CALEPSCGCTPLP 417

Db 905 KSMVNAETIRKHLPOYEDQIRKITINALPLQDSSLVWLPLVSKSGEYTTKTGVALAKLNSFPAS 964

QY 418 SM-----ASTRAARLGELLQ-----DPRSHQCOQWMDNNA 449
Db 965 QLENNQKINIKIHTSPKVNHFILMKAMKALPYGEALSRNIEAVTCRCGO--TESSL 1022
QY 450 GSIACPMKDYIHRRWVYLWVLAACLLAALFEFLKKDRKARGSRALLHSADGAG 509
Db 1023 HMLLCP---YAKKVE-----LAPVLF-----NPSEATHSSVALLLVDA---- 1059
QY 510 YERLVGALSALSOMPLR--VAVDLMSRRELGAHGAALAMFHQRRIIOEGGVILLFSP 567
Db 1060 -KRMVALPPTGASAPLYPWALMHEH-----KARNRLIFDNHS-----CSEGLVLF----- 1104
QY 568 AAVAOCCOMLQLOTV-----EPCGPHAL-----AAMLS--CVLPDFLO---- 603
Db 1105 KAILDARAMEADLLIHPSPISDYPSPTPLIKYTSCEFVDAAWTTSGYCGMGWFLQDPYK 1164
QY 604 -----GRATGRYVG 612
Db 1165 VKIKENSSSSPYG 1178

RESULT 15

T26240
Hypothetical protein W06F12.1a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26240
R:Lloyd, C.
Submitted to the EMBL Data Library, December 1996
A:Reference number: Z20179
A:Accession: T26240
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-647 <MIL>
A:Cross-references: EMBL:Z83244; PIDN:CAB05827.1; GSPDB:GN00021; CESP:W06F12.1a
A:Experimental source: clone W06F12
C:Genetics:
A:Gene: CESP:W06F12.1a
A:Map position: 3
A:Introns: 47/2; 87/1; 145/1; 179/1; 216/3; 255/2; 317/2; 353/1; 564/3; 634/2

Query Match 2.7%; Score 100.5; DB 2; Length 647;

Best Local Similarity 23.0%; Pred. No. 6.7;

Matches 45; Conservative 30; Mismatches 72; Indels 49; Gaps 8;

QY 85 RVVVLAVHGHAEPEEAG-----KSDSELQSRNLSLAQVVLSPQAVPIARCALLEV 138
Db 176 QLVGEMALVSH--THRAVSTTCYEKNQKQOQVQOIPQOV-----AHVSSNAIILAA 229
QY 139 Q-----VPADLVQPGQSVGSAVDFCEASLGAEOIWSYTKRYOKEL-----NLTOOLP 188
Db 230 QPFYPPYQDSQPDRIIGYGAGCV-----VMSVTDPKSGKRVALKKMPNPFQNL 279
QY 189 DC-----RGLQVDSIOSCWVLPRLANSTGDNVLLTLDVSEEDFSFLLYLRPPDALKS 244
Db 280 SCKRVFRETIKMLSSFRH-----DNVLSLIDILQPNPSPFOELVYVTELMQS 326
QY 245 LMYKNLTGPNITLNH 260
Db 327 DLKRTIVSPQALTPDH 342

Search completed: September 28, 2002, 19:41:32
Job time: 6105 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

protein - protein search, using sw model

on: September 28, 2002, 19:40:27 ; Search time 27.72 Seconds

(without alignments)
615,046 Million cell updates/sec

File: US-09-899-471-5
Effect score: 3741
Sequence: 1 MPVSFLLSLALGRNPVVS.....SSAPGCGCEMDIGPCTTLE 698

oring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

arched: 231628 segs, 24425594 residues

*total number of hits satisfying chosen parameters: 231628

imum DB seq length: 0
ximum DB seq length: 200000000

st-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued Patents, AA.*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Hit No.	Score	Query Match	Length	ID	Description
1	312	8.3	617	4	US-09-188-930-303
2	177	4.7	864	2	US-08-620-694A-2
3	177	4.7	864	3	US-09-022-255-2
4	177	4.7	864	3	US-09-022-255-2
5	177	4.7	864	3	US-08-978-773-2
6	177	4.7	864	3	US-09-022-253-2
7	177	4.7	864	3	US-09-022-253-2
8	177	4.7	864	4	US-09-022-259-2
9	177	4.7	864	4	US-09-022-257-2
10	150.5	4.0	330	4	US-09-188-930-125
11	147.5	3.9	866	2	US-08-620-694A-10
12	147.5	3.9	866	3	US-09-022-255-10
13	147.5	3.9	866	3	US-09-022-255-10
14	147.5	3.9	866	3	US-08-978-773-4
15	147.5	3.9	866	3	US-09-022-253-10
16	147.5	3.9	866	3	US-09-022-260-10
17	147.5	3.9	866	4	US-09-022-259-10
18	147.5	3.9	866	4	US-09-022-257-10
19	112	3.0	1447	5	US-09-041-886-25
20	112	3.0	1447	5	PCT-US94-05277-2
21	107.5	2.9	1711	2	US-08-342-930-2
22	103	2.8	1410	4	US-09-335-409-3
23	103	2.8	1410	4	US-09-568-102-3
24	103	2.8	1410	4	US-09-567-969-3
25	103	2.8	1410	4	US-09-568-480-3
26	103	2.8	1410	4	US-09-568-486-3
27	103	2.8	1410	4	US-09-568-472-3

28	97.5	2.6	1618	1	US-07-853-913-4	Sequence 4, Appl
29	95.5	2.6	1498	2	US-08-404-531B-28	Sequence 28, Appl
30	95.5	2.6	1498	3	US-08-476-900A-28	Sequence 28, Appl
31	95.5	2.6	1498	3	US-08-488-546A-28	Sequence 28, Appl
32	95.5	2.6	1581	3	US-08-404-531B-6	Sequence 6, Appl
33	95.5	2.6	1581	3	US-08-476-900A-6	Sequence 6, Appl
34	95.5	2.6	1581	3	US-08-488-546A-6	Sequence 6, Appl
35	95.5	2.6	2629	2	US-08-751-189-4	Sequence 4, Appl
36	95.5	2.6	2629	2	US-09-060-836-4	Sequence 4, Appl
37	95.5	2.6	2629	2	US-09-184-445-4	Sequence 4, Appl
38	94.5	2.5	2627	2	US-08-751-189-3	Sequence 3, Appl
39	94.5	2.5	2627	2	US-09-060-836-3	Sequence 3, Appl
40	94.5	2.5	2627	4	US-09-184-445-3	Sequence 3, Appl
41	93.5	2.5	726	4	US-09-126-980-2	Sequence 2, Appl
42	93.5	2.5	726	4	US-09-476-482-2	Sequence 2, Appl
43	93.5	2.5	880	1	US-08-445-640-10	Sequence 10, Appl
44	93.5	2.5	880	3	US-08-170-558-10	Sequence 10, Appl
45	93.5	2.5	880	3	US-08-447-314-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-188-930-303
; Sequence 303; Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Marison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188, 930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FASTSeq for Windows Version 3.0
; SEQ ID NO 303
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Mouse
US-09-188-930-303

Query Match	8.3%	Score 312;	DB 4;	Length 617;
Best Local Similarity	23.9%	Pred. No. 1.4e-24;	Indels 136;	Gaps 21;
Matches 134;	Conservative	66;	Mismatches 224;	
OY	175	PRYQKELINTQQLDPCRGLEY-----RDSIQSC--WVLPMILNVSFGDGNVLLTIDVSEE	226	
DB	137	PERSFPL-----LPEQAVAVRTIPAGPKASVRLCYQWALFEEDLSSPDT-----	181	
OY	227	QDESFLLLRPVVDALKSLMYKNLTGPONTILNHTDLPFCIGIOWSLSPDSEVEECPF	286	
DB	182	-----OKIVSGGHTVDLP-----LPLPCMCIEASYIOEDIVRRKKCFE	221	
OY	287	REDPGAH-RULMHIARLRVLSFG-----WMDAPCCLREKTKLQWAPQOSQPLVPPFP	342	
DB	222	QSMPELVGSDFMOSIRFTDYSOHNOMVALLRCPKLKIASLCWRQDDPLTPPETL-----	276	
OY	343	QKNATYNEQDFQLVAG---HPNLCYQVSTWEKYQLOAC-----SWADSLGPFKDM	391	
DB	277	-FNATQSEGWYILENVLDHPQLCFKFS-FENSSHCEPCQSGSLPMTYSMDT-QAQO	333	
OY	392	LVEAKTGTGLNNTSVCLERSGC---TPLPSMASTRARLGEELLQD-----FRSHQCMQL	443	
DB	334	LTLHSSRTYATFSAAWSPDGLCPDTPMPVYSISQTOGSAVPTLDLITFLRGNCILY	393	
OY	444	WMDNNGLSLMACPMKXTHRRWLV-----WLACILLAAALFFELLKKDRRAAGS	496	

Db 394 WRSD-----VHFARKHVLCPDDAPYPTOLL-----BSLGSRT 427

Qy 497 FTALLHSADAGYERLVGALASALSQM---PLRVAVDLMSRRELNAHGLAMFHHQRR 553

Db 428 RPVLLHHADEAQRRLVGALELLTALGGRDVLVDMEGTHVAIGPLPMMAARER 487

Qy 554 ILORGCVILLFSPRAVACQOWIOLQYEPFPHDLAAMLSCVLPDFLOGRATGRYGV 613

Db 488 VAROGVYLLLMNCGPGS-----TACSGDPQOASLRTLLCAAPRPL-----LLA 531

Qy 614 YFDLLHPDSVPSPRVAPFLFELPTOLPAFLDALOGGCSFSA-----GRPADREVERVQ 667

Db 532 YFSNLCAKAGDIPRPLRALPRVLLDLPLRLALDAPATLASSWSHLGAKRCLKNRLQ 591

Qy 668 A-----LRSAUDSCYSSSEAP 683

Db 592 CHLELEAKKDYGSTNSP 611

RESULT 2

US-08-620-694A-2

Sequence 2, Application US/08620694A

Patent No. 5869286

GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin

APPLICANT: Spriggs, Melanie

APPLICANT: Fanslow, William

TITLE OF INVENTION: No. 5869286el Receptor That Blinds IL-17

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple Power Macintosh

OPERATING SYSTEM: Apple Operating System 7.5.5

SOFTWARE: Microsoft Word for Apple, Version 6.0.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/620,694A

FILING DATE: 21 MARCH 1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/538,765

FILING DATE: 7 AUGUST 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/410,535

FILING DATE: 23 MARCH 1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,695

REFERENCE/DOCKET NUMBER: 2617-B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)587-0430

TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 864 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-620-694A-2

Query Match 4.7%; Score 177; DB 2; Length 864;

Best Local Similarity 20.5%; Pred. No. 1.1e-09;

Matches 121; Conservative 65; Mismatches 189; Indels 216; Gaps 29;

Qy 198 SIQSCW-----VLPMINVTSGDNDVLLTLDVSEODEFSFLYLRYPV----- 239

Db 2 AIRCCEPRVYVPGALGWL-----LLLNVLAPGRASPRLLDPFPAVCAQEGISCRV 52

Qy 240 ---DALKSLWY---KNLT---GQNTITLN-----HTDLVPCLCIOVWSLEPD----- 277

Db 53 KNSCTLDDSWIHPKNTLPPSSPKNIYINLVSSTQGEGLVPLVHVE-WTLOTASIIYLKG 111

Qy 278 -----SER---VECPREDPGARHNLHTRLR-VLSFG-----VQDLAPCCL 318

Db 112 AELSVLQLTNTERLCVAF-OFLSMLOHHRKWRFSFSHFVVDGGEYEVYVHLLPRP--- 167

Qy 319 PGKVTLCQWAPDQSPQCP-----LVPPVPOKNATVNEPODFOLVAGHPNLQVGVSTWEKVQ 374

Db 168 -----IPDGPNNKSKIIIFVPCDEDSKMKMTSCVSSGSLMDPNITVETLDTQHLR 218

Qy 375 LQACSWADSLGPFERDMLYEMKTGLNNTSV----- 405

Db 219 VDFTLNNEST-PLYO-----VLSEFSDSEHNSCFDVKQIFAPROEFHORANTFTLSKPH 274

Qy 406 -CA-----LEP---SGC-----TPLPMSASTRAARLSEGLLODFRSHQCMQIMND 447

Db 275 WCGHHVQVOPFFSSCLNDCLRHAVTVPQPVISMTTVPK----- 313

Qy 448 NMGSIMACPMKXYIHRKV--LVWLACLLLAALFFFL-----LKKDRRKARGS--- 496

Db 314 -----PVADVI-PLMYGLIILAILLVGSYVILLICMTWRISGADQERHGDSTKN 364

Qy 497 -----FTALLHSADGAGYERLVGALASAL-SQMDPLRAVDLMSRRELSA 540

Db 365 GLIPVADLTPEPLRPKRWIYVSADHPLYEVVLKFAQFLTYACGTEVALDLLEQVISE 424

Qy 541 HGALWPFHHORRRHIOGCVIILFS-----PAAVACQOWIOLQYVEEG 585

Db 425 VGVMTWVSROKQEWESNKTIIILCSRGTOAKKAILGLMAEPVAVDLRCDHW-----KPA 478

Qy 586 PHDALAAMLSCVLPDFLOGRATGRYGVYFEDGLHPDSVPSPFRAVPLFSL 636

Db 479 -GDLFPAAMMILPDRFRACFGYVVCYFSGICSERVDPOLFNTSRPL 528

RESULT 3

US-09-022-255-2

Sequence 2, Application US/09022255

Patent No. 6072033

GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin

APPLICANT: Spriggs, Melanie

APPLICANT: Fanslow, William

TITLE OF INVENTION: No. 6072033el Receptor That Blinds IL-17

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple Power Macintosh

OPERATING SYSTEM: Apple Operating System 7.5.5

SOFTWARE: Microsoft Word for Apple, Version 6.0.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/022,255

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/620,694

FILING DATE: 21 MARCH 1996

APPLICATION NUMBER: USSN 08/538,765

FILING DATE: 7 AUGUST 1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/410,535
 FILING DATE: 23 MARCH 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia Anne
 REGISTRATION NUMBER: 34,695
 REFERENCE/DOCKET NUMBER: 2617-B
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206)587-0430
 TELEFAX: (206)
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 864 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-022-255-2

Query Match 4.7%; Score 177; DB 3; Length 864;
 Best Local Similarity 20.5%; Pred. No. 1.1e-09;
 Matches 121; Conservative 65; Mismatches 189; Indels 216; Gaps 29;

198 SIQSCW-----VLPWLNVSTGDDNVLLTLDVSEEDDFSLYLIRPV----- 239
 2 AIRCWPVRVPGPALGWL-----LILLNLVAPGRASPRLLDFPAPVCAQEGLSGRV 52
 240 --DAKSLWY--KNLT--GPNITLN-----HTDLVPCICIOVWSLEPD----- 277
 53 KNSTCLDSDWIHPKMLTPSPKNIYINLSSSTOIHGELVPLVHVE-WTLOTDA SILYLEG 111
 278 -----SER--VECPREDPGAHNLMHIAIRL-VLSPG-----VWOLAPCCL 318
 112 AELSVLQNTNERLCVKE-QFLSMLOHHRKRRFSFHFVDDPGQEXEVYVHHLPRK--- 167
 319 PGKVTLCQWAPDQSPQCP---LVPPVPOKNATVNEPDDQVLVAHGNLVCQVSTWEKVO 374
 168 -----IPDCDPRNHSKIIFVPCDEDSKMKMTTSCVSSGSLMDPNITVETLDTQHILR 218
 375 LQACSWADSLGPRKDDMLVEMKTGLNNTSV----- 405
 219 VDFTLNNEST-PLYO---VLLESFSDSENHSCFDVYVKOIFAPROEEFHORANVTFLSKFH 274
 406 -CA-----LEP--SGC-----TPLPMSASTRAARLGEELLDFRSHQCMQIMND 447
 275 WCCHHHVOVQPFSSCLNDCLRHAVTVPCEVISTNTVYK----- 313
 448 NMSGSLACPMKTYIHRNV--LVWLACLLLAALFEFL-----LKKDRRKAARG-- 496
 314 -----PVADYI-PLWVYGLITLIALIVGSVIYLIICMTWRLSGADQEKHGDSKIN 364
 497 -----RTALLHSADGAGYERLVGALASAL--SQMPLRVANVDMRRELSA 540
 365 GILPVADITLPPRLPRKRWIYISADHPLEYVVLKRAQFLITACGVEVALDLLEQVISE 424
 541 HGALAMFHQRRILOEGGVYLLFS-----PAVAACQOQWMLDQUTVEPG 585
 425 VGVATWVSROKQEVESNKIILICSRGQAKMKAILGMAEPVQALCDHW-----KPA 478
 586 PHDLAMLSVLDLDFLOGKATGRVGYTFDGLLHSDVSPSPRVAPLFSL 636
 479 -GDLETAAMNMLPDKFRPACFGTYVCFSGICSERDVPDLFNITSRYPL 528

RESULT 4

US-09-022-696-2

Sequence 2, Application US/09022696

Patent No. 6072037

GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin

APPLICANT: Spriggs, Melanie

APPLICANT: Fanslow, William

TITLE OF INVENTION: No. 6072037a1 Receptor That Binds IL-17

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Immunex Corporation
 STREET: 51 University Street
 City: Seattle
 STATE: WA
 COUNTRY: USA
 ZIP: 98101
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: Apple Power Macintosh
 SOFTWARE: Microsoft Word for Apple, Version 6.0.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/022,696
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/620,694
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/620,694
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia Anne
 REGISTRATION NUMBER: 34,695
 REFERENCE/DOCKET NUMBER: 2617-B
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206)587-0430
 TELEFAX: (206)
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 864 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-022-696-2

Query Match 4.7%; Score 177; DB 3; Length 864;
 Best Local Similarity 20.5%; Pred. No. 1.1e-09;
 Matches 121; Conservative 65; Mismatches 189; Indels 216; Gaps 29;

198 SIQSCW-----VLPWLNVSTGDDNVLLTLDVSEEDDFSLYLIRPV----- 239
 2 AIRCWPVRVPGPALGWL-----LILLNLVAPGRASPRLLDFPAPVCAQEGLSGRV 52
 240 --DAKSLWY--KNLT--GPNITLN-----HTDLVPCICIOVWSLEPD----- 277
 53 KNSTCLDSDWIHPKMLTPSPKNIYINLSSSTOIHGELVPLVHVE-WTLOTDA SILYLEG 111
 278 -----SER--VECPREDPGAHNLMHIAIRL-VLSPG-----VWOLAPCCL 318
 112 AELSVLQNTNERLCVKE-QFLSMLOHHRKRRFSFHFVDDPGQEXEVYVHHLPRK--- 167
 319 PGKVTLCQWAPDQSPQCP---LVPPVPOKNATVNEPDDQVLVAHGNLVCQVSTWEKVO 374
 168 -----IPDCDPRNHSKIIFVPCDEDSKMKMTTSCVSSGSLMDPNITVETLDTQHILR 218
 375 LQACSWADSLGPRKDDMLVEMKTGLNNTSV----- 405
 219 VDFTLNNEST-PLYO---VLLESFSDSENHSCFDVYVKOIFAPROEEFHORANVTFLSKFH 274
 406 -CA-----LEP--SGC-----TPLPMSASTRAARLGEELLDFRSHQCMQIMND 447
 275 WCCHHHVOVQPFSSCLNDCLRHAVTVPCEVISTNTVYK----- 313
 448 NMSGSLACPMKTYIHRNV--LVWLACLLLAALFEFL-----LKKDRRKAARG-- 496
 314 -----PVADYI-PLWVYGLITLIALIVGSVIYLIICMTWRLSGADQEKHGDSKIN 364
 497 -----RTALLHSADGAGYERLVGALASAL--SQMPLRVANVDMRRELSA 540

Db 365 GILPVADLPPLPRKRWIYVSADHPLVEVLAKEAFLITACGTEVALDLLEOVISE 424
QY 541 HGALAMFHQRRILOEGVILLES-----PAVAOCOQWLOLQTEVPG 585
Db 425 VGVMTWVSROKOEWEENSKIIILCSRGTOAKKAILGMAEPAVOLRCDHW-----KPA 478
QY 586 PHDALAMLSVLPDLQGRATGRYGVYFDGLLHPDSVPSFRRVAPLESL 636
Db 479 -GDLETAAMMILPDKFRPACFTYVVCYFSGICSERDVPDLFNITSRYPL 528
RESULT 5
US-08-978-773-2
Sequence 2, Application US/08978773
Patent No. 6083906
GENERAL INFORMATION:
APPLICANT: Trout, Anthony
TITLE OF INVENTION: Method of Regulating Nitric Oxide Production
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple PowerMacintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for PowerMacintosh, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,773
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/052,525
FILING DATE: 27 NOVEMBER 1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2623-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 864 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-978-773-2
Query Match 4.7%; Score 177; DB 3; Length 864;
Best Local Similarity 20.5%; Pred. No. 1.1e-09;
Matches 121; Conservative 65; Mismatches 189; Indels 216; Gaps 29;
QY 198 SIOSCM-----VLPMLNSTDGNVLLTLDVSEQDSFLLYLRVP-----239
Db 2 AIRCWMRVVPGPALGWL-----LLLLNVLAAGRASPRLLDPAPCAQEGSLSCR 52
QY 240 ---DALKSLMY--KNLT--GPNITLN-----HTDLVPCLCIQWMSLEPD-----277
Db 53 KNSICLDSDSHHPNLTSPSSKNNTIINLSVSTQHGELVPLAVE-WTLQTDASITYLEG 111
QY 278 -----SER--VECCPREDDPGAHRLNMLHIALRL-VLSPG-----VMQIDAPCCL 318
Db 112 AELSLVQLNTNERLCVKE-QELSMLOHHRKRRFSEFSHVVYDPGQGEYTVHLLPR---167
QY 319 PGKVTYLCWQAPDSQPCP-----LVPPVPQKNAVTNEPQDFQLVAGHPNLGVQVSTWEKYQ 374

Db 168 -----IPDGPNHKSKITFPDCEDSKMKKTTSCVSSGSLMDPNTIVETLDTQHLR 218
QY 375 IQACSNADSLGPEKDMILVEMKGLNNTSY-----405
Db 219 VDTFLMNEST-PIQ---VLIESFSDSENHSCPDVYKQIPAPROEFHORANVTFLSKFH 274
QY 406 -CA-----LEP--SGC-----TPLPSMASTRARLAGEELLQDFRSHOCQMLMND 447
Db 275 WCCHHHVQVPFESSCLNDCRHAHTVPCPIYSNTYK-----313
QY 448 NMGSLNACPMCKYIHRKV--LVMLACLLLAALFFPL-----LLKDRRAARGSS---496
Db 314 -----PVADYI-PLMYVGLITLIALILVGSYVLIICMTWRLSGADQEKHGDDSKIN 364
QY 497 -----RRALLHSADGAGYERLNGALASL--SOMLRVAVDMSRRELSA 540
Db 365 GILPVADLPPLPRKRWIYVSADHPLVEVLAKEAFLITACGTEVALDLLEOVISE 424
QY 541 HGALAMFHQRRILOEGVILLES-----PAVAOCOQWLOLQTEVPG 585
Db 425 VGVMTWVSROKOEWEENSKIIILCSRGTOAKKAILGMAEPAVOLRCDHW-----KPA 478
QY 586 PHDALAMLSVLPDLQGRATGRYGVYFDGLLHPDSVPSFRRVAPLESL 636
Db 479 -GDLETAAMMILPDKFRPACFTYVVCYFSGICSERDVPDLFNITSRYPL 528
RESULT 6
US-09-022-253-2
Sequence 2, Application US/09022253
Patent No. 6096305
GENERAL INFORMATION:
APPLICANT: Yao, Zhengbin
APPLICANT: Spriggs, Melanie
TITLE OF INVENTION: No. 6096305el Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,253
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/620,694
FILING DATE: 21-MARCH-1996
APPLICATION NUMBER: USSN 08/538,765
FILING DATE: 7 AUGUST 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

Sun Sep 29 09:32:51 2002

us-09-899-471-5.raii

Page 5

```

; LENGTH: 864 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
DS-09-022-253-2

```

Query Match	4.7%;	Score 177;	DB 3;	Length 864;
Best Local Similarity	20.5%;	Pred. No. 1.1e-09;		
Matches 121;	Conservative 65;	Mismatches 189;	Indels 216;	Gaps 29

QY	198	SIOGCM-----	VLPLYNASTODGNVLLTLDSEEDQFSLYLARVP-----	239
DB	2	AIRRCMRRVVRGRLGMD-----	LLLLLVNLABCRBRASRLDDPAPRCADQGLSCRY	52
QY	240	---DALKSMTW--KNLT--GPNITLN-----	HTDVLPCICLOVMSLEPD-----	277
DB	53	KNSTCLDDSMWHPRKNLPRSPKNIYILNLSVSTQHGLVPLVANE--	WTLOQDASILYLEG	1111
QY	278	---SER--VEFCPRREDPGAHRLIMHTARL--VLSFG-----	VMOLADPCL	318
DB	112	AELSVLOLNTNERLCYAF--QFLSMQDHHRRKMRFSFHFVDDPQOEYEYVHNLKRP---	167	
QY	319	PGKVTLCMAQAPDQSPCQR-----	LYPRPQKNATVNEQDQDLVAGHNLLOVYSTWCKVQ	374
DB	168	-----IPDGDPRNKSXKIIIPVPCDEDSKMKWTJSCVSSGLMDNITVETLDTQHLR	218	
QY	375	LQACSMADSLGPRFMDLILVEMKTGLNNTSV-----	405	
DB	219	VDFLTNNEST--PRQ--VLLSEFSDSNSHSCFDVVKQIFAPROEFPQORANVTYLSKFH	274	
QY	406	-CA--LEP--SGC-----	TPRPSMASTRARLGEELLDFRSHOCQOLMND	447
DB	275	WCCHHHVOVOPRFSSCINDCLRHAUTVPCFVISTNTVPR-----	313	
QY	448	NMGSLMCPDKTYTHRMV--LWMLACLLLAALFFL-----	LKKDRKKAARG--	496
DB	314	-----PVADYI--PLMYTGLITLAILLVGSAIVLIICTMRRLSGADEKGGDSKIN	364	
QY	497	-----RTALLHSADGAGYERLVGALASAL--SOMPLRVAVDLMSRRELSA	540	
DB	365	GILPVADLPRPLPRKRVWVLTYSADHNLIEYVULKRAQFLITACGEVVALDLLEBOYISE	424	
QY	541	HGALAMFHQRRILODEGVVILIFS-----	PAVAOQCOOMLOLQTYEPG	585
DB	425	VGVMTVWSRQKQEWENESKIIILCSKGTQAKKATLIGMAEPVQJLRCDHM-----	KPA	478
QY	586	PHALAAAMTSCVLPDLOGKATRGARYGVYPRDGLLHPDSVSPRPVAPLPSL	636	
DB	479	-GDLFTRAAMNLTEDFRKPRACFGTYVVCYFSGICSEBDYVPLENTISRRPL	528	

RE: RESULT 7
US-09-022-260-2
Sequence 2, Application US/09022260
Patent No. 6100235
GENERAL INFORMATION:
APPLICANT: Yao, Zhengbin
APPLICANT: Spriggs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6100235e1 Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5

```

SOFTWARE: Microsoft Word for Apple, Version 6.0.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,260
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/620,694
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 864 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-260-12

```

Query Match	4.7%	Score 177	DB 3	Length 864
Best Local Similarity	20.5%	Pred. No. 1.1e-09		
Matches 121	Conservative 65	Mismatches 189	Indels 216	Gaps 29

QY	198	SIO\$CW-----VLPMLN\$TDBDANVLITLDVSEODF\$SLYLARVP-----	239
Db'	2	AIRRCMPRVVGPALGWL-----LLLLNLVAIPERASRILLDPAPPCAOGLSCRV	52
QY	240	---DAKSLW--KNLT--GPONITL-----HTDLVPLCLIQWSLEBD-----	277
Db	53	KNSTCILDSDSWIHRLNPSSPKNIYIMLSYSTOIGELVPLHVE-WTLQTDSILYLEG	111
QY	278	-----SER--VECEPFREDDGAHRMLMHLARK-ULSPG-----WMOLADPCL	318
Db	112	AELSVDLOLNTNERLCYAF-OFLSMLQHRRRWRESFSHFAYVDPGEYEVTVHLRP-	167
QY	319	PGRYTLCOWARDSPFCOP-----LYVRPOKNAUYNEPDQOLVAGHNLCUYOSTWEKV	374
Db	168	-----IPDGDPRNKSKIIFPPDDEDSKMKMTTSCVSSGSLMDPNITYETILOHKL	218
QY	375	LQACSNADSI\$GPFKKDMLVEMKTGLNITSV-----	405
Db	219	VDELINME\$T-PYO-----VLLESFSDSENHC\$FDVKOIFAPROEEHQANVTFTLSKFH	274
QY	406	-CA-----LBP--SGC-----TPUSMA\$T\$BARAGEELLDFRSHOCQOLMND	447
Db	275	WCCHHHVOVOPFPSSCINDCLRHAATVPCEVISNTVPK-----	313
QY	448	NMSL\$MACPDKTYIHR\$MV--LWJLACLALAAE\$PFL-----LLKDRRKARG\$--	496
Db	314	-----PADYI-PLMYGLITLAILLVGSYIVLIICMTWRLSGAODEKHGGDSKIN	366
QY	497	-----RTALLHSADGAGERYLVGALASAL-SOMPLRYAVADLMSRRELTA	540
Db	365	GILPVADLP\$PLRPRKVMYIVYSADHPLYEVVLKFAQOLITACGTBEVALADLLEQYISE	424
QY	541	HGALANFHNRKRITLDEGVYIILFS-----PAAVAQCQOMILOQTVEPG	565
Db	425	VGMVTVWSRKOEBVESNSKIITLIC\$RGTOAKKAIIGNAEP\$VOLRCDHM-----KPA	478
QY	586	PHNALAMTSCLVEDLOGRATGRYGVYUYPDGILHDPVS\$P\$RVAPL\$FSL	636
Db	479	-GLETR\$ANM\$ILLDERKRAPCFGYVVYVCSGCISERDVDELNITRSRYPL	528

gaps
of
the
structure
of
the
system

55 SAGP-----VLVP-----TRQTEVLELRCPQKTCALRRVVVHILAVHGHNPE 100
: ||| ||| ||| ||| ||| : |||


```

10 AVPGPLGLLLLLGLVLAAGSASLRLLDHRALVCSOP---GLNCTVKNSTCLDDSWIHR 66
101 EAGKSDSELOESRNASIAQAVLVSF---QAYPIARCALLEVOYPADLVPOGSGVSAVF 156
67 -----NLTPSSPKDQIOHLFHAHTQGGDLFPVAH---IEMTLQTD----- 103
157 DCFEASL-----GAEOVIMSTKPRYOKELNLTQQLPDCRGLEVRDSIQ---SCWVLPMLN 209
104 -----ASITYLEGAEISV-----LQNTNERL---CVREFLSKLRHHRRRFFESH 148
210 VSTDGDNVLLTLVDSEEDFSFLYLRLVPDALKSLMYKNLTGPONITLNNHLDVPCLC- 268
149 FVVDPD-----QEYEVYVHHLKPRIPDGDPRNOSKNFLVPCDHAKMKVTPPCMS 199
269 -----IQWLSLEPDSERVEFC-----PREDPGAHNMLMHIARLVLS 306
200 GSLMDPNITVETLEAHQLRVSFTLMNESTHYOILLTSPHMHNSCFEHHNHI----- 252
307 PGVWQDAPCCLPKGYTLQWQARDSPCOPLYRPPVPOKNAIVNEPDFO-----LVAGH 360
253 -----PAPR-----PEEFHORSNVTLTLRN 272
361 PNLCVOVSTWEKVOLOACSWADSLGPFKDDMLLVEMKTGLNNTSVCLSPSGCTPLPSMA 420
273 LKGGCR---HOVOIQ-----PF-----FSSCLND---CLRHSAIVSCPEMP 307
421 STRAARLGEELLQDFRSHQCMQMLMNDNMGSLMACPMKYIHRWVLYWLACLLLAALF 480
308 DT-----PEPIPDY-----MPLM-----YWFITGISILLVGSVI- 337
481 FELLKKDRRKARGS-----RALLHSADGAGYERLVG 515
338 -LLIVCMTWRLAGPSEKYSDDTKYTDGLPAADLIPPLKPRKWIIYSADHPLVVDVVL 396
516 ALAS-ALSOMPLRAVADLMSRRRLSAHGLAMFHHORRILQEGGVYLLFSPAANAQO 574
397 KRAQFLITACGEVALDLEQAISEAGVMTVGROKQEWESKTIIVLCSRGTRAKWQ 456
575 QWL-----QLOTVEPGP-HDALAAMLSCVLPDLQGRATGRVGVYFDGILLHDSVSP 627
457 ALLGRGAPVRLRCHGKVGDLFTAAAMNILLDFKRPACGFTGVVGYEFSVSCGDVBDL 516
628 FRVAPFSLPQLDPAF-----LDALOGCSTSAGRAPRAREVVO-----A 668
517 FGAAFRPLMDREFEYFRIDLEMFOPG-----RMHRVGLSGDNYLRSRPGRO 566
669 LRSALD 674
567 LRAALD 572

```

SEQUENCE 12

US-09-022-255-10

Sequence 10, Application us/09022255

Patent No. 6072033

GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin

APPLICANT: Spriggs, Melanie

APPLICANT: Fanslow, William

TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: Apple Power Macintosh

SOFTWARE: Microsoft Word for Apple, Version 6.0.1

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,255
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USN 08/620,694
FILING DATE: 21 MARCH 1996
APPLICATION NUMBER: USN 08/538,765
FILING DATE: 7 AUGUST 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USN 08/410,535
FILING DATE: 23 MARCH 1995
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 866 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-255-10

```

Query Match 3.98; Score 147.5; DB 3; Length 866;

Best Local Similarity 19.3%; Pred. No. 1.7e-06;

Matches 140; Conservative 78; Mismatches 239; Indels 269; Gaps 32;

```

55 SAVGP-----YVVP-----TRLOTELYRCPQKTDCAIRVVVYVHVAHGHAPE 100
10 AVPGPLGLLLLLGLVLAAGSASLRLLDHRALVCSOP---GLNCTVKNSTCLDDSWIHR 66
101 EAGKSDSELOESRNASIAQAVLVSF---QAYPIARCALLEVOYPADLVPOGSGVSAVF 156
67 -----NLTPSSPKDQIOHLFHAHTQGGDLFPVAH---IEMTLQTD----- 103
157 DCFEASL-----GAEOVIMSTKPRYOKELNLTQQLPDCRGLEVRDSIQ---SCWVLPMLN 209
104 -----ASITYLEGAEISV-----LQNTNERL---CVREFLSKLRHHRRRFFESH 148
210 VSTDGDNVLLTLVDSEEDFSFLYLRLVPDALKSLMYKNLTGPONITLNNHLDVPCLC- 268
149 FVVDPD-----QEYEVYVHHLKPRIPDGDPRNOSKNFLVPCDHAKMKVTPPCMS 199
269 -----IQWLSLEPDSERVEFC-----PREDPGAHNMLMHIARLVLS 306
200 GSLMDPNITVETLEAHQLRVSFTLMNESTHYOILLTSPHMHNSCFEHHNHI----- 252
307 PGVWQDAPCCLPKGYTLQWQARDSPCOPLYRPPVPOKNAIVNEPDFO-----LVAGH 360
253 -----PAPR-----PEEFHORSNVTLTLRN 272
361 PNLCVOVSTWEKVOLOACSWADSLGPFKDDMLLVEMKTGLNNTSVCLSPSGCTPLPSMA 420
273 LKGGCR---HOVOIQ-----PF-----FSSCLND---CLRHSAIVSCPEMP 307
421 STRAARLGEELLQDFRSHQCMQMLMNDNMGSLMACPMKYIHRWVLYWLACLLLAALF 480
308 DT-----PEPIPDY-----MPLM-----YWFITGISILLVGSVI- 337
481 FELLKKDRRKARGS-----RALLHSADGAGYERLVG 515
338 -LLIVCMTWRLAGPSEKYSDDTKYTDGLPAADLIPPLKPRKWIIYSADHPLVVDVVL 396
516 ALAS-ALSOMPLRAVADLMSRRRLSAHGLAMFHHORRILQEGGVYLLFSPAANAQO 574
397 KRAQFLITACGEVALDLEQAISEAGVMTVGROKQEWESKTIIVLCSRGTRAKWQ 456
575 QWL-----QLOTVEPGP-HDALAAMLSCVLPDLQGRATGRVGVYFDGILLHDSVSP 627

```


Db 457 ALIGGAPVRLCDHGRKPVGDLFTFAAMNMLPDKRPACGTYVVCSEVSCDGDVDDL 516
Qy 628 FRVAPLESLPTQLPAF-----IDLALOGCSTSGRPRADVERVTO-----A 668
Db 517 FGAPRYPLMDREFEYVFRIDOLEMFQPG-----RMRVGLSGDNYLSPGGRQ 566
Qy 669 LRSALD 674
Db 567 LRAALD 572

RESULT 13
US-09-022-696-10
Sequence 10, Application US/09022696
GENERAL INFORMATION:
APPLICANT: Yao, Zhenobin
APPLICANT: Spriggs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6072037e1 Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,696
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/620,694
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995

ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 866 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-696-10

Query Match 3.9%; Score 147.5; DB 3; Length 866;
Best Local Similarity 19.3%; Pred. No. 1.7e-06;
Matches 140; Conservative 78; Mismatches 239; Indels 269; Gaps 32;

Qy 55 SARP-----VLVP-----TRLOTELVLRCPKTDKALRVVVVHLAVHGAEP 100
Db 10 AVGPPLGLLLLLGLVLAFGASLRILDRALVCSQ-----GLNCYVKKSTGCLDDSMIHR 66
Qy 101 EAGKSSSELEQESNALSQAVLSF-----QATYPARCALLEVQVADLVQSGVSAVF 156
Db 67 -----NLTPSSPKDILQIOLAFHATQGGDLFPVAH-----IEMTLOTD----- 103

Qy 157 DCEPASY-----GAEOVIMSYTKPRYOKELNLTQOLPDCRGLEVRDSIQ---SCWVLPWIN 209
Db 104 -----ASTLYEGAEIASV-----LQNTNERL---CVREFELSKLRHHRMRFTESH 148
Qy 210 VSTDGQNVLLTLDVSEBODESEFLLYLRVPDPALSKLWTKNLTPGONITLNTHTDLVPCIC- 268
Db 149 FVVDPD-----QREYTVHHLKPKIPDGDPRNHSKNFLVPCELAKMVTTPCMS 199
Qy 269 -----IQWLSLEPDSERVEFC-----PEREDPGAHNLMIARLYLS 306
Db 200 GSLMDPNITVEETLEAHLQVLRVSFTLMNESTHYQILLTSPHEMNHSCFEMHNI----- 252
Qy 307 PGWQOLDAPCCLPGKVTILCMQAPDQSPQPLVPPYPOKNAIVNEQDQ-----LVAGH 360
Db 253 -----PAPR-----PEEFHORSNVTLTLRN 272
Qy 361 PNLGVSTWERYOLQACSWADSLGPEKDDMLLVEMKGTGLNNTSVCALEPSCGTFPLPSMA 420
Db 273 LKGCGR-----HQVQIQ-----FSSCLND-----CLRHSATVSCPEMP 307
Qy 421 STRAARLGEELLODFRSHQCMQIMNDNMGSLMACPMOKYIHRKVLVWLACLLAALF 480
Db 308 DT-----PEPIDY-----MPLM-----VWFITGISILVGSVY- 337
Qy 481 FFLLEKKDRKKAARG-----RTALLHSADGAGYERLVG 515
Db 338 -LITVCMTRNLAPGSEKYSDDTKYTDGLPRAADLIIPPLKRWIITYSADHPIVVDVL 396
Qy 516 ALAS-ALSQMPLRAVAVDLMRSRELTAHGLALWAFHRRRILOEGGVVILFSPAVAOCQ 574
Db 397 KPAQFLTAQGTVALDLEQAISEGVTWVGRQDQEMYESKTIYLSKRGTRAKMQ 456
Qy 575 QNL-----QLQTVRGP- HDALAAMLSCVLPDLQGRATRGYGVYFGDLHPDSVSP 627
Db 457 ALIGGAPVRLCDHGRKPVGDLFTFAAMNMLPDKRPACGTYVVCSEVSCDGDVDDL 516
Qy 628 FRVAPLESLPTQLPAF-----IDLALOGCSTSGRPRADVERVTO-----A 668
Db 517 FGAPRYPLMDREFEYVFRIDOLEMFQPG-----RMRVGLSGDNYLSPGGRQ 566
Qy 669 LRSALD 674
Db 567 LRAALD 572

RESULT 14
US-08-978-773-4
Sequence 4, Application US/08978773
Patent No. 6083906
GENERAL INFORMATION:
APPLICANT: Trout, Anthony
TITLE OF INVENTION: Method of Regulating Nitric Oxide Production
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple PowerMacintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for PowerMacintosh, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,773
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/052,525
FILING DATE: 27 NOVEMBER 1996

CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia Anne
 REGISTRATION NUMBER: 34,693
 REFERENCE/DOCKET NUMBER: 2623-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206)587-0430
 TELEFAX: (206)
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 866 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 OS-08-978-773-4

Query Match 3.9%; Score 147.5; DB 3; Length 866;
 Best Local Similarity 19.3%; Pred. No. 1.7e-06;
 Matches 140; Conservative 78; Mismatches 239; Indels 269; Gaps 32;

55 SAPP-----VLVP-----TRLOTETVLRCPKQKTDCAALRVRVVLAHGHAEPE 100
 10 AVPGPLGLILLGLLVLPAGGASRLDLHRLVCSQP---GLNCTVKNSTCLDDSWIHR 66
 101 EAGKSDSELOESRNASIAQAVVLSF---QAYPIARCALLEVQVPADLVQPGQSVGSAVF 156
 67 -----NLTPSSPKDQLQIHLFAHTQGGDLFPVAH---IEWTLQTD----- 103
 157 DCFEASL-----GAEOVIMSYTKPRYQKELNLTQOLPDCRGLEVRISIQ---SCWVLPWLN 209
 104 -----ASILLEGAEISV-----LQNTNERL--CVREFELSKLRHHRRMRTFSH 148
 210 VSTDGDNVLLTLDVSEEDFSEFLYLRLPVDPALKSIMYKNLTGPONITLNIHDLVPCIC- 268
 149 FVVDPD-----DEYEVTVHILPKPIPDGDPNHQSKNPLVPCENARMKVTPFCMS 199
 269 -----IOWMSLEPDSEREREC-----PREDDPCAHNMLHIAIRLVLS 306
 200 GSLMDPNITVETLEAQLRVSTFLMNSTHYQILLTSPHMHNSCFEIMHI----- 252
 307 PCWQMDAACCLCGKVTYLCQAPQSPCQPLVPVPQKNATVNEPDQF-----LVAGH 360
 253 -----PAPR-----PEERHONSNVTLTLNN 272
 361 PHLGVQVSTWEKVOQLQACSWADSLGPPKDMLLVEMKTGLNNTSVCALEPSCGTEPLPSMA 420
 273 LRGCCR---HQVOIQ-----PF-----FSSCLND---CLRHSATVSCPEMP 307
 421 STRAARLGEELLQDFRSHQCMQMLMNDNMGSLMACPMADKYIHRRVVLYWLACLLLAALP 480
 308 DT-----PEPIPDY-----MPLM-----VYWFITGISILVLGSVI- 337
 481 PFLLKKDRRKARGS-----RTMLLLSADGAGTERLVG 515
 338 -LLIVCMTRIRLAPGSEKSYDDTKYTDGLRAADILPPPLKFRKRWIITISADHPLVYDVL 396
 516 ALAS-ALSCMPLVAVDLMSRRELISAHGALAMFHNRRLIQEGGVVILSPAAVQCO 574
 397 KPAOELITACGTEVALDLLEEQAISEAGVMTVGRQKQEMVESNKIIVLCSRGRAKQ 456
 575 QWL-----OLQTVGERP-HDALAAMLSVLPDLOGRATRYGVYFDGLHPDSVSP 627
 457 ALLGKCAPVRLKCDHKRPVGDLEFTAANMMLLPDKRPACFCGTYYVVCYSEVSCDDVDPL 516
 628 FRVAPLFSPTQLPAF-----LDALOGGCSSTSAGRPAADRYERTQO-----A 668
 517 FCAAPRYPLMDREERYFYRIQDLEMFORG-----RMRHVGEISGDVNLSPGGRQ 566
 669 LRSAID 674
 567 LRRAID 572

RESULT 15
 US-09-022-253-10
 : Sequence 10, Application us/09022253
 : Patent No. 6096305
 : GENERAL INFORMATION:
 : APPLICANT: Yao, Zhengbin
 : APPLICANT: Spriggs, Melanie
 : APPLICANT: Fanslow, William
 : TITLE OF INVENTION: No. 6096305el Receptor That Binds IL-17
 : NUMBER OF SEQUENCES: 10
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Immunex Corporation
 : STREET: 51 University Street
 : CITY: Seattle
 : STATE: WA
 : COUNTRY: USA
 : ZIP: 98101
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: Apple Power Macintosh
 : OPERATING SYSTEM: Apple Operating System 7.5.5
 : SOFTWARE: Microsoft Word for Apple, Version 6.0.1
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/022,253
 : FILING DATE:
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/08/620,694
 : FILING DATE: 21-MARCH-1996
 : APPLICATION NUMBER: USSN 08/538,765
 : FILING DATE: 7 AUGUST 1995
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: USSN 08/410,535
 : FILING DATE: 23 MARCH 1995
 : CLASSIFICATION:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Perkins, Patricia Anne
 : REGISTRATION NUMBER: 34,695
 : REFERENCE/DOCKET NUMBER: 2617-B
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (206)587-0430
 : TELEFAX: (206)
 : INFORMATION FOR SEQ ID NO: 10:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 866 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-09-022-253-10

Query Match 3.9%; Score 147.5; DB 3; Length 866;
 Best Local Similarity 19.3%; Pred. No. 1.7e-06;
 Matches 140; Conservative 78; Mismatches 239; Indels 269; Gaps 32;

55 SAPP-----VLVP-----TRLOTETVLRCPKQKTDCAALRVRVVLAHGHAEPE 100
 10 AVPGPLGLILLGLLVLPAGGASRLDLHRLVCSQP---GLNCTVKNSTCLDDSWIHR 66
 101 EAGKSDSELOESRNASIAQAVVLSF---QAYPIARCALLEVQVPADLVQPGQSVGSAVF 156
 67 -----NLTPSSPKDQLQIHLFAHTQGGDLFPVAH---IEWTLQTD----- 103
 157 DCFEASL-----GAEOVIMSYTKPRYQKELNLTQOLPDCRGLEVRISIQ---SCWVLPWLN 209
 104 -----ASILLEGAEISV-----LQNTNERL--CVREFELSKLRHHRRMRTFSH 148
 210 VSTDGDNVLLTLDVSEEDFSEFLYLRLPVDPALKSIMYKNLTGPONITLNIHDLVPCIC- 268
 149 FVVDPD-----DEYEVTVHILPKPIPDGDPNHQSKNPLVPCENARMKVTPFCMS 199

Sun. Sep. 29 09:32:51 2002

us-09-899-471-5.rai

Page 11

```

01 269 -----IOWSLERPSRVERFC-----PFREDPGARNMLWHTARLVLS 306
02 200 GSIIMDPNIVETTELHAHOIRSFETLMNESTHYOILTFSPHMENHSCFEHNHHT----- 252
03 307 PGWQIDAPCCLPCKVTLCMOARQDQSPQPLVPBPVQKATYNEPDQO-----LYAGH 360
04 253 -----PAPR-----PEEFORSNWTLTLRN 272
05 361 PNICVOYSTWEKVOLAQCSWADSLGPFKDDMLVEMKGTANTSVCALEPSCGTPLPBSMA 420
06 273 LKGCGR--HQVQIQ-----PF-----FSSCLND-----CLHRSATVSCPEMP 307
07 421 STPAAKLGEELODFRSHQCMOLNDDNMGSLMACPMKYTHRRVLYWVACILLLAAALF 480
08 308 DT-----PEPIPDY-----MPLW-----VYMFITGISILVGSVI- 337
09 481 FFLTKKRRKRAARGS-----RTALLHSADGAGYERLVG 515
10 338 -LIVICMTWMLAGBSKYSDDTYTQGLPRADLIRPPLKPRKWIYYSADHPILYDVVL 396
11 516 ALAS -ALSQMPLRAVVDLMSBRELISAGCALAMEHHORRRRIIOEGGVILLFSPAANAQO 574
12 397 KRFQFLITACTEVALDILLEQOAISEAGVMTWVGROKQEWESNSKIITVLCGRGTRAKQO 456
13 575 QWL-----OLQTVBERP- HDALAAWLSCVLPDFLOGRATGRYGVGFEDGJLHPSVSP 627
14 457 ALLGGRAPVLRCDHGKRPVDLLFTFAANNMILPDKRRACPGTYVVCYFSEVSCDDGVPDL 516
15 628 FRVAPLFSLETPQLPAF-----LDALOGGCSAGRAPRVERVTO-----A 668
16 517 FGAAPRYPLMDREVEEYFRIODLEMFQPG-----RMHRVGLSGDNVLRSPGGRQ 566
17 669 LRSALD 674
18 567 LRRALD 572

```

Search completed: September 28, 2002, 19:40:30
Job time: 6273 sec

Of: US-09-899-471-2 to: N_Geneseq_032802:* out_format : pfs
Date: Sep 29, 2002 1:21 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame.p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09899471/runat_27092002.142440.17943/app-query.fasta.1.2519
-DB=N_Geneseq_032802 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000
-GAPOP=4.500 -GAPPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FCGAPOP=6.000 -FCGAPPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blonsum62
-TRANS=human0.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pts
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09899471.@CGN1_1_926 -NCPU=6 -ICPU=3 -LONGIOG
-DEV_TIMECUT=120 -WARN_TIMECUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-899-471-2
Query length: 674
Database: N_Geneseq_032802:*
Database sequences: 1736436
Database length: 858457221
Search time (sec): 974.100000

Score list:

Sequence	Strid Orig	ZScore	EScore	Len	Documentation
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AA518132 +		3565.00	4661.71	2.6e-251	23
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AA518133 +		2637.00	3444.34	1.7e-183	20
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA55027 +		2352.00	3070.09	1.2e-162	2
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA56223 +		2339.00	3051.83	1.2e-161	2
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA509515 +		2339.00	3051.83	1.2e-161	2
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA52138 +		2339.00	3051.83	1.2e-161	2
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AA518130 +		2235.00	2915.60	4.7e-154	23
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA55029 +		2182.00	2845.23	3.9e-150	2
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA55029 +		1851.50	2414.84	3.7e-126	1
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA55028 +		1671.00	2177.87	5.8e-113	17
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AA55028 +		1652.50	2151.90	1.6e-111	2
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AA518131 +		1548.00	2014.54	7.3e-104	21
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA55030 +		1291.50	1679.79	3.2e-85	1
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA59805 +		843.50	1095.14	1.2e-52	1
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA52817 +		692.50	870.94	3.6e-40	1
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA57657 +		692.50	868.31	5.1e-40	2
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA578822 +		670.00	869.45	4.4e-40	2
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA54469 +		604.50	788.26	1.5e-35	6
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA548278 +		599.50	780.20	4.1e-35	7
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA540777 +		552.00	717.96	1.2e-31	6
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA542818 +		367.00	459.20	3.1e-17	3
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA52815 +		361.00	451.21	8.7e-17	3
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AA518136 +		328.00	413.30	1.1e-14	20
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA509516 +		316.50	397.93	8.5e-14	2
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA599796 +		310.50	389.63	2.4e-13	4
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA561753 +		310.50	389.59	2.4e-13	21
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA569666 +		310.50	389.59	2.4e-13	2
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA54216 +		309.50	400.81	5.6e-14	4
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA572373 +		291.00	364.10	6.2e-12	4
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA54821 +		230.00	289.05	9.4e-08	1
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA52815 +		205.00	269.39	1.2e-06	3
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA518131 +		181.50	222.29	0.0005	1
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:AA573800 +		174.50	206.89	0.0036	32
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AA527591 +		174.50	206.89	0.0036	32
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AA501921 +		174.50	206.89	0.0036	32
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA561238 +		174.50	206.89	0.0036	32
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA51987 +		174.50	206.89	0.0036	32
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA52145 +		174.50	206.89	0.0036	32
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA59870 +		174.50	206.89	0.0036	32
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA502813 +		174.50	206.89	0.0036	3

/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA57186 + 174.50 206.89 0.0036
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:AA573801 + 154.00 180.17 0.1092
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AA527592 + 154.00 180.17 0.1092
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AA501922 + 154.00 180.17 0.1092
/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA561240 + 154.00 180.17 0.1092

seq_name: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AA518132

seq_documentation_block:

ID AA518132 standard; cDNA: 2314 BP.
AA518132;
26-MAR-2002 (first entry)
Mouse DNAX cytokine receptor subunit 7 (DCRS7) cDNA.
Mouse; DNAX cytokine receptor subunit 7; DCRS7: phosphate labelling; ss;
gene therapy; protein therapy; immunological disorder.
Mus musculus.
Key CDS Location/Qualifiers
CDS 199..2295
FT /**tag=a
/product="Mouse DCRS7"
MO200190358-A2.
29-NOV-2001.
23-MAY-2001: 2001MO-US16767.
24-MAY-2000: 2000US-206862P.
(SCHE) SCHERING CORP.
Gorman DW;
WPI: 2002-106198/14.
P-PSDB; AA011354.
Isolated antigenic human or mouse DNAX receptor subunit-1like
polypeptide useful for detecting antibodies generated in response to
presence of increased protein levels or immunological disorders -
Disclosure; Page 17-20; 148pp; English.

The invention relates to primate and rodent DNAX cytokine receptor
subunit (DCRS) polypeptides and the polynucleotides encoding them. The
receptors, or their portions may be useful as phosphate labelling enzymes
to label immunogens or specific substrates. The subunits may also be
functional immunogens to elicit recognising antibodies, or antigens
capable of binding antibodies. A combination, e.g., including a DCRS can
be used as an immunogen for the production of antisera or antibodies
capable of distinguishing between other cytokine receptor family members.
A purified DCRS can also be used as a reagent to detect antibodies
generated in response to the presence of elevated levels of expression,
or immunological disorders which lead to antibody production to the
endogenous receptor. This sequence represents cDNA encoding the mouse
DCRS7 polypeptide.

Sequence 2314 BP; 411 A; 725 C; 705 G; 473 T; 0 other;

alignment_scores:

Quality: 3565.00 Length: 698
Ratio: 5.305 Gaps: 1
Percent Similarity: 96.275 Percent Identity: 96.132

alignment_block:

US-09-899-471-2 x AA518132

Align seg 1/1 to: AAS18132 from: 1 to: 2314

1	MetProValSerTrpPheLeuLeuSerLeuAlaLeuGlyArgAsnProVa	17
199	ATGGCTGTGGTCCCTGGTTCCTGGCTCCTTGCGCACTGGGCCAAACCTGT	248
17	IValValSerLeuGluArgLeuMetGluProGlnAspThrIleArgCysS	34
249	GCTGCTCTCTCTGGAGACACTGATGGAGCTCCAGGACACTCACCTCT	298
34	erLeuGlyLeuSerCysHisLeuTrpAspLysAspValLeuCysLeuPro	50
299	CTTAGAGGCTCTCCAGCCACCTCTGGGATGGTGAAGCTGTGCCTGCT	348
51	GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuGl	67
349	GGAAAGCCCTCCACTGCCCCAGGCCCTGTGCTAGTGCCTACCCTGCA	398
67	nThrgLuLeuValLeuArgCysProGlnLysTrpAspCysAlaLeuArg	84
399	GACGAGCTGTGCTGAGGTGCTCCACAGAAGACAGATTGCGCCCTGTG	448
84	aLArgValValValHisLeuAlaValHisGlyHisTrpAlaGluProGlu	100
449	TCCGTTGTGTGCTCCACTGGCCGTGCATGGCAGCTGGCGAGACCTTA	498
101	GluAlaGlyLysSerAspSerGluLeuGlnGluSerArgAsnAlaSerLe	117
499	GAACCTGGAAAGTCTGATTCCAGAACCTCCAGAGACTCTAGGAAGCCTCT	548
117	uGlnAlaGlnValValLeuSerPheGlnAlaTrpProIleAlaArgCysA	134
549	CCAGGCCAGGAGGTGGTCTCTCCAGGCTCACCCCATGCCCCTGTG	598
134	IaLeuLeuGluValAlaGlnValProAlaAspLeuValGlnProGlyGlnSer	150
599	CCCGTGGAGAGTCCAGAGTCCCGCGTGCAGCTGGTGCAGCCTGGTCACTGC	648
151	ValGlySerAlaValAlaPheAspCysPheGluAlaSerLeuGlyAlaGluVa	167
649	GTGGGTCTCGGGTATTGACTGTTTCGAGAGCTAGTCTTGGGGCTGAGCT	698
167	LgIleIleTrpSerTrpThrLysProArgTrpGlyIndylsGluLeuAsnLeut	184
699	ACAATCTGGGTCTACACAGAAACCCAGGTACACGAAGAAGCTCAACCTCA	748
184	hrrGlnGlnLeuPro.....	188
749	CACAGCAGCTCGCTACAGCGGCTTGGAATCCGGGACAGCATCCAG	798
189 AspLysAspAsnVa	193
799	AGCTGTGGGTCTCCCTCGGCTCATGTGTCTACAGATGGTGCATATGT	848
193	IleuLeuThrLeuAspValSerGluGluGlnAspPheSerPheLeuLeut	210
849	CCCTTGTACACTGTGATGTCCTGAGAGAGCAGGACTTACGCTTCTTACTGT	898
210	yrlLeuArgProValProAspAlaLeuLysSerLeuTrpTryLysAsnLeu	226
899	ACCTGCGCTCCAGTCCCGGATGGCTCTCAAAATCTTGGTGCACAAAACCTG	948
227	ThrGlyProGlnAsnIleThrLeuAsnHisThrAspLeuValProCysLe	243
949	ACTGAGACTCGAACAATTACTTAAACACACAGACTGGTTCCTGCT	998
243	wCysIleGlnValTrpSerLeuGluProAspSerGlyArgValGluPheC	260
999	CTGCATTCCAGGTGTGGTGTGCTAGACCAACACTGTGAGAGGGTGCATTC	1048
260	ysrProPheArgGluAspProGlyAlaHisArgAsnLeuTrpHisIleAla	276
1049	GCCCCCTTCGGGAAGATCCCGGTGCACACAGGAACCTTGGCACATGACC	1098

277	ArguanaAgValLeuSerProGlyValITripGlnLeuAspAlaProCys	293
1099	AGCGTCGGGTAATCTCTCCAGAGGATATGGACACTGATGTCGGCTTCCTG	1148
293	sLeuProGlyValThLeuCysTrpGlnAlaProAspGlnSerProC	310
1149	TCTGGCGGGCAAGTAAACACTGTGTCTGGCAGACCAACAGATCTCT	1198
310	ysGlnProLeuValProProValProGlnLysAsnAlaThrValAsnGlu	326
1199	GCCAGCCCACTGTGTGCCACAGTGGCCCCACAAAGAACCCACATGTGAATGAG	1248
327	ProGlnAspPheGlnLeuValAlaGlyIleProAsnLeuysValGlnVa	343
1249	CCACAAAGATTTTCAGTTGGTGGCAGCCACCCCAACCTGTGTGCCAGGT	1298
343	IserThrTrpGlnLysValGlnLeuGlnAlaCysSerTyrAlaAspSerL	360
1299	GAGCACTCGGGAGAGAGTTCACACTGCCAACGTCCTGTGGCGTACTCT	1348
360	eugLyProPheLysAspAspMetLeuValGlnMetLysThrGlyLeu	376
1349	TGGGCGCCCTTCAAGATGATTAATGCTGTATGTGAGATGAAGAAACCGGCTC	1398
377	AsnAsnThrSerValCysAlaLeuGluProSerGILCysThrProLeuPr	393
1399	AACAACACATCATGTCGTGTGCTTGGAAACCCAGTGGCTGTACACACTGCC	1448
393	oSerMetAlaSerThrArgAlaAlaArgLeuGlyGlnLeuLeuGlnA	410
1449	CAGATGGCCCTCCACAGAGCTGCTCCCTCGGGAGAGAGTGTCTGCCAG	1498
410	sPheArgSerHisGlnCysMetGlnLeuTyrAsnAspAspAsnMetGly	426
1499	ACTTCGATTCACACCAAGTATATGCAGCTTGGACCATGACAACTAGGA	1548
427	SerLeuTyrAlaCysProMetAspLysTyrIleHisArgTyrTrpValLe	443
1549	TCCGTATGGGCTGGCCCATGACAAATGATCATCCACAGCGGCTGGGCTCT	1598
443	uValTyrPheuAlaCysLeuLeuLeuAlaAlaAlaLeuPhePheLeuL	460
1599	AGTATGGCTGGCTGCTCTACTTGTGGCTGGCGGCTTTCCTTCTCTCC	1648
460	euleuLysLysAspArgTyrGlyAlaAlaArgGlySerArgThrAlaLeu	476
1649	TTCTAAATAAAGACCGCAGAAAGCGGCCGCGTCCGACAGGCGCTTG	1698
477	LeuLeuHisSerAlaAspGlyAlaGlyTyrGlnArgLeuValGlyAlaLe	493
1699	CTCTCTCCACTCCGCGAGCGAGCGGGCTATACGACGCGCTGTGGAGACT	1748
493	uAlaSerAlaLeuSerGlnMetProLeuArgValAlaValAspLeuTrpS	510
1749	GGCGTCCGCGCTTGACCGACAGATGCCACTGCGCTGGCCGTAAGACTTGGA	1798
510	eArgArgGlnLeuSerAlaHisGlyAlaLeuAlaLysPheHisHisGln	526
1799	GCCGCGCCGAGACTGAGCGCGACGAGGCCCTGTAGCCGTTCACACACAG	1848
527	ArgArgArgGlyLeuGlnGlyGlyValValIleLeuLeuPheSerPr	543
1849	CGAGCGCGGTATCTTCGACGAGAGGTGGCTGTGAATCTTCTTCTTCGCC	1898
543	oAlaAlaValAlaGlnCysGlnGlnTyrProLeuGlnIleuGlnThrValGluP	560
1899	CGCGCGCGCTGGCGGAGTGTACGACATGGCTGTGACGTCCAGACAGTGGACC	1948
560	roGlyProHisAspAlaLeuAlaAlaTyrProLeuSerCysValLeuProAsp	576
1949	CCGGGCGCGAAGAGCCCTTCGCCGCTGTGCTAGCTGTGCTATCCCGAT	1998

-name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAS18133

```

- documentation_block:
  AAS18133 standard; cDNA; 2094 BP.

```

AAS18133-

26-MAR-2002 (first entry)

Mouse DCRS7 reverse translation generic cDNA

mouse; DNAX cytokine receptor subunit 7; DCRS7; phosphate labelling; ss gene therapy; protein therapy; immunological disorder.

Mus musculus.

W0200190358-A2.

29-NOV-2001.

23-MAY-2001; 2001WO-US16767.

24-MAY-2000; 2000US-206862P.

(SCHE) SCHERING CORP.

Gorman DM;

WPI: 2002-106198/14.

Disclosure: Page 20-21; 146pp; English.

The invention relates to primate and rodent DNA cytokine receptor subunit (DCRS) polypeptides and the polynucleotides encoding them. The receptors, or their portions may be useful as phosphate labelling enzymes to label general or specific substrates. The subunits may also be functional immunogens to elicit recognising antibodies, or antigens capable of binding antibodies. A combination, e.g., including a DCRS can be used as an immunogen for the production of antisera or antibodies capable of distinguishing between other cytokine receptor family members. A purified DCRS can also be used as a reagent to detect antibodies generated in response to the presence of elevated levels of expression, or immunological disorders which lead to antibody production to the endogenous receptor. This sequence represents mouse DCRS7 reverse translation generic cDNA.

XX	Sequence
SQ	2094 BP; 247 A; 251 C; 393 G; 285 T; 918 other;

alignment_scores:

Quality: 2637.00	Length: 698
Ratio: 4.938	Gaps: 1
Percent Similarity: 76.504	Percent Identity: 69.198

alignment_block:
US-09-899-471-2 x AAS18133

Align seg 1/1 to: AAS18133 from: 1 to: 2094

[illegible]

	seq_name: /SIDS1/gcdata/geneseq/geneseg-emb1/NA2001A.DAT.AAC85027
seq_documentation_block:	
ID AAC85027 standard; cDNA; 2255 BP.	
XX AC AAC85027;	
XX DT 08-MAY-2001 (first entry)	
XX DE Human cytokine receptor zcytor14 encoding cDNA.	
XX KW Cytokine receptor; zcytor14; human; inflammatory; rheumatoid arthritis;	
XX KM antiinflammatory; gene therapy; vaccine; ss.	
XX OS Homo sapiens.	
XX FH Key Location/Qualifiers	
FT CDS 3..2232	
FT FT /*tag= a	
XX XX /product= "Zcytor14"	
PN WO200104304-A1.	
PD 18-JAN-2001.	
PX 30-JUN-2000; 2000WO-US18383.	
PR PF	
PR 07-JUL-1999; 99US-0348854.	
XX PA (ZYMO) ZYMOGENETICS INC.	
PI Presnell SR, Burkhead SK, Powder SL;	
DR WPI: 2001-112618/12.	
DR P-PDB: AAB61880.	
XX New polypeptide encoding a human cytokine receptor zcytor14, for	


```

471  lySerArgThrAlaLeuLeuHisSerAlaSpGlyAlaGlyTyrGlu 487
      ::::::::::::::::::::::::::::::::::::::::::::
1601  GGGGGCCCCGGGCTGTGCTCTCTACTAGCCGAGACGCGGCTTCGAG 1650
      ArgLeuValGlyAlaLeuAlaSerAlaLeuSerGlnMetProLeuArg 504
      ::::::::::::::::::::::::::::::::::::::::::::
1651  CGGCTGTGGGGCCCGCTGGCTCGGCTGTGCGAGCTCCCTGCGCGT 1700
      504  AlaValAspLeuTrpSerArgArgGluLeuSerAlaHisGlyAlaLeu 521
      ::::::::::::::::::::::::::::::::::::::::::::
1701  GGCGGTGACCTGTGAGCGCTCGTAGACTGAGCGGCGGCGCGTGG 1750
      521  lAtrrPheHisHisGlnArgArgArgGluGluGluGlyAlaVal 537
      ::::::::::::::::::::::::::::::::::::::::::::
1751  CTGCGTTTCACGCGCAGCGCGGCGCAGACCTGCGAGAGGCGCGGTG 1800
      538  lLeuLeuPheSerProAlaAlaValAlaGlnCysGlnGlnTrpLeu 554
      ::::::::::::::::::::::::::::::::::::::::::::
1801  GTCTTGTCTTCTCTCCGGGTGGGTGGCTGTGACGAGGTGCTTACA 1850
      554  nLeuGlnThrValGluPro.....GlyProHisAspAlaLeuAla 568
      ::::::::::::::::::::::::::::::::::::::::::::
1851  GGATGGGCTGTCCGGGCGCGGCGGCGGCGGCGGCGGCTTCGCGG 1900
      568  lAtrrPheSerCysValLeuProAspPheLeuGlnGlyArgAlaThrGly 584
      ::::::::::::::::::::::::::::::::::::::::::::
1901  CCTCGCTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1950
      585  ArgTyrValGlyAlaValTyrPheAspGlyLeuLeuHisProAspSerVal 601
      ::::::::::::::::::::::::::::::::::::::::::::
1951  AGCTAGCTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2000
      601  oSerProPheArgValAlaProLeuPheSerLeuProThrGlnLeuPro 618
      ::::::::::::::::::::::::::::::::::::::::::::
2001  CGCCCTTTTCGCGACCGCTGCGCTTACACTGCTCCCGCAGCTGCGG 2050
      618  lAphLeuAspAlaLeuGlnGlyCysSerThrSerAlaGlyArgPro 634
      ::::::::::::::::::::::::::::::::::::::::::::
2051  ACTTCTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
      635  AlaAspArgValGluArgValThrGlnAlaLeuArgSerAlaLeuAsp 651
      ::::::::::::::::::::::::::::::::::::::::::::
2101  CAGAAGAGAGCGGAGCAGAGTGTCCCGGCTTCCAGCCAGCCCTGATG 2150
      651  rCys.....ThrSerSerGlnAlaProGly 660
      ::::::::::::::::::::::::::::::::::::::::::::
2151  CTACTTCATCCCGCGGAGCTCCCGCGCGGGA 2184
      CT-name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DMT:AAS46223
      documentatn_block:
      AAS46223 standard: CDNA: 2380 BP.
      AAS46223:
      18-DEC-2001 (first entry)
      Human DNA encoding PRO polypeptide sequence #299.
      PRO polypeptide: mammal: tumour: cancer: human: cattle: horse: sheep: ss:
      dog: cat: pig: goat: rabbit: tumour necrosis factor alpha; TNF-alpha;
      blood: chondrocyte cell; cell proliferation; cell differentiation; colon;
      adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
      PCR primer.
      Homo sapiens.
      WO200168848-A2.
      20-SEP-2001.
      28-FEB-2001; 2001WO-US06520.

```

```

XX 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 03-MAR-2000; 2000US-187202P.
PR 06-MAR-2000; 2000US-186968P.
PR 14-MAR-2000; 2000US-189320P.
PR 14-MAR-2000; 2000US-189328P.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000US-190828P.
PR 21-MAR-2000; 2000US-191007P.
PR 21-MAR-2000; 2000US-191048P.
PR 21-MAR-2000; 2000US-191314P.
PR 28-MAR-2000; 2000US-192655P.
PR 29-MAR-2000; 2000US-193032P.
PR 29-MAR-2000; 2000US-193053P.
PR 30-MAR-2000; 2000WO-US08439.
PR 04-APR-2000; 2000US-194449P.
PR 04-APR-2000; 2000US-194647P.
PR 11-APR-2000; 2000US-195975P.
PR 11-APR-2000; 2000US-196000P.
PR 11-APR-2000; 2000US-196187P.
PR 11-APR-2000; 2000US-196690P.
PR 11-APR-2000; 2000US-196820P.
PR 18-APR-2000; 2000US-198121P.
PR 18-APR-2000; 2000US-198585P.
PR 25-APR-2000; 2000US-199397P.
PR 25-APR-2000; 2000US-199550P.
PR 25-APR-2000; 2000US-199654P.
PR 03-MAY-2000; 2000US-201516P.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 05-JUN-2000; 2000US-209832P.
PR 28-JUL-2000; 2000WO-US20710.
PR 22-AUG-2000; 2000US-064848.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US4956.
XX (GETH ) GENENTECH INC.
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2001-602746/68.
XX P-PSDB; AA029322.
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
XX presence of tumours, such as prostate and breast tumours, in mammals and
XX to screen for modulators of the compounds -
XX Claim 2; Fig 597; 774pp; English.
XX Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR
XX primers for PRO polypeptides of the invention. The sequences of the
XX invention can be used to detect the presence of a tumour in a mammal by
XX comparing the level of expression of a PRO polypeptide in a test sample
XX of cells from the animal and a control sample of normal cells, whereby a
XX higher level of expression in the test sample indicates the presence of a
XX tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,
XX pigs, goats and rabbits but are preferably human. The polypeptides can
XX be used to stimulate tumour necrosis factor (TNF) alpha release from human
XX blood, when contacted with it. A specific polypeptide can be used to
XX stimulate the proliferation or differentiation of chondrocyte cells. The
XX PRO proteins can be used to determine the presence of tumours and also
XX susceptibility to tumour development, particularly adrenal, lung, colon,
XX breast, prostate, cervical, or liver tumours, in mammalian
XX subjects. The oligonucleotide probes specific for the PRO nucleic acids
XX can be used for genetic analysis of individuals with genetic disorders.
XX Sequence 2380 BP; 411 A; 776 C; 743 G; 450 T; 0 other;

```


PR 02-MAR-2000; 2000OMO-US05841.
PR 21-MAR-2000; 2000US-0191007.
PR 21-MAR-2000; 2000OMO-US07532.
PR 02-JUN-2000; 2000OMO-US15264.
PR 22-JUN-2000; 2000US-0213087.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000OMO-US23328.
PR 24-OCT-2000; 2000US-0242837.
PR 10-NOV-2000; 2000OMO-US30873.
PR 28-NOV-2000; 2000US-0253646.
PR 01-DEC-2000; 2000OMO-US32678.

XX
XX PA (GENENTECH INC.)
PI PI Chen J., Flvaroff E., Fong S., Goddard A., Godowski P.J., Grimaldi C.T.,
PI Gurney AL., Li H., Hillan KJ., Tunas D., Van Lookeren M., Vandlen RL,
PI Watande CK, Williams PM, Wood WI, Yansura DG;
XX WPI: 2001-451708/48.
DR P-PSDB; AAU04956.

PT Novel PRO polypeptides homologous to interleukin-17, useful for the
PT diagnosis and treatment of immune related disease e.g. Rheumatoid
PT arthritis and diabetes -

XX
XX PS Claim 1; Fig 13; 188pp; English.

CC The sequence (DNA 164625-2890) encodes a PRO polypeptide (PRO20040)
CC which is the human Interleukin 17 receptor, IL-17RH. A composition
CC containing ant/agonists to the PRO polypeptides or individual components
CC are useful for treating a mammal with an immune related disease, e.g.
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease, an autoimmune or immune-mediated skin
CC disease, contact dermatitis, an allergic disease e.g. food
CC hypersensitivity, asthma, a transplantation associated disease, or a
CC chronic inflammatory demyelinating polyneuropathy. Treating a
CC degenerative cartilaginous disorder comprises administering a PRO1031 or
CC PRO1122 polypeptide agonist, or antagonist to the mammal. Numerous
CC examples of the diseases and disorders are given in the specification.

XX
XX SQ Sequence 2380 BP; 411 A; 776 C; 743 G; 450 T; 0 other;

Alignment_scores:
Quality: 2339.00 Length: 692
Ratio: 4.118 Gaps: 7
Percent Similarity: 82.081 Percent Identity: 66.329

alignment_block:
US-09-899-471-2 x AAS09515 ..

Align seg 1/1 to: AAS09515 from: 1 to: 2380.

1 MetProValSerTrpPheLeuLeuSerLeuAlaLeuGlyValAsnProVal 17
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
233 AGCGCGTGCCTCGTTCTTGCTGCTCCTTGCGACTGGCGGACCAACCAGT 282
17 |ValValSerLeuGluArgLeuMetGluProGlnAspThrAlaArgCysS 34
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
283 GGTCTTCCTTCGTGGAGAGCGCTTGCGGCCCTCAGGACGTACCCACATGCT 332
34 etLeuGlyLeuSerCysHisLeuTrpAspGlyAspValIleuCysLeuPro 50
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
333 CTCGGGCGCTTCCTCCTCCGCGCTTGCGACAGTACATCATCTGCTGCT 382
51 GlySerLeuGlnSerAlaProGlyProValIleValProThrArgLeuG 67
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
383 GGGAACAATGCTGCTCCGGGCGCGCTGCTGACGCTACGACACATCGA 432

1333	CTTGGGGCTGACTCCCTGGGGCTTCCAAAGACAGATGTGCTATTGGG	1387
371	lumeLlysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluProSer	387
1383	AGACACGAGGCCCCAGGACAAACAGATCCCTCTGTGGCTTGGAAACCA	1433
388	GlyuysThrProLeuProSerMetAlaSerThrArgAlaAlaArgLeuG	404
1433	GGCTGTACTTCACTCCCAAGCAAAGCCTCCACAGAGGACAGTCCCTGG	1482
404	UGlUGlUleuLeuGlnAspPheArgSerHisGlnCysMetGlnLeuP	421
1483	AGAGTACTTACTACAAAGACCTGCAGTCAGGCCCCAGTGTCTGACGTAT	1531
421	snAspAspAsnMetGlySerLeuThrAlaCysProMetAlaPlyuTyrl	437
1532	..GAGAGTACTTGGGAGGCGATATGGGCGCTGCCCAATGACAAATACAT	1579
438	HisArgArgThrValLeuValThrPheAlaCysLeuLeuAlaAlaAl	454
1580	CACAAGCCTGGGCGCTGTGTGGCTGGCGCTTCACTCTTGGCGGCTGC	1629
454	AlaPhePhePheLeuLeuLeuLysAspArgAlaGlyLys.....	467
1630	GGTTTCCCTCATCTCTCTTCAAAAAAGATACACCGAAAGGGTGGCTGA	1679
468AlaAlaArgGlySerArg	473
1680	GGCTTTGAACAGACAGCTCCCTGGGGGCGGCGCCCAAGGGG...CGC	1726
474	ThrAlaLeuLeuLeuHisSerAlaAspArgAlaGlyGlyGlyAlaVal	490
1727	CGGCTGTGCTCTGTACTACAGCCGATGACTGGGTTTGCAGGGCGCTGG	1776
490	IGlYAlaLeuAlaSerAlaLeuSerGlnMetProLeuArgValAlaVal	507
1777	GGGGCGCTGGCGCTGGCGCTGGCACAGCTGGCGCTGGCGGTGGCGCTAG	1826
507	spLeuTrpSerArgArgGlyLeuSerAlaHisGlyAlaLeuAlaTrpPhe	523
1827	ACCTGTGGAGCGGTGTACTACAGCGCGCACGGAGCGCGCGCTTGGTTT	1876
524	HisHisGlnArgArgArgLLeuGlnGlnGlyValValAlaLeuLeu	540
1877	CAGCGCAGCGCGCGCACACCTGCAGAGAGGGCGGCTGGTGGTCTGGCT	1926
540	upheSerProAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnLeuGln	557
1927	CTTCTCTCCCGGTGCGGCGGCGCTGCAGAGAGTGGTACAGAGAGGGG	1976
557	hValGluPro.....GlyProHisAspAlaLeuAlaAlaTrpLeu	570
1977	TGTCGGGGCGGGGGCGACGAGCGCGCACAGCGCGCTGGCGCGCTGCT	2026
571	SerCysValLeuProAspPheLeuGlnGlyArgAlaHisGlyArgGly	587
2027	AGCTGGCTGCTGGCGGACTTTTGCAGGGCGCGGGCGCGGACAGTCTG	2076
587	IGlValuTrpPheAspArgLeuLeuHisProAspSerValProSerPro	604
2077	GGGGGCGCTGCTGCACAGGTGTCTCCACCGGAGCGGAGCGGACCGCTTT	2126
604	hArgValAlaProLeuPheSerLeuProGlnGlnLeuProAlaPheLeu	620
2127	TTCGCAACGCGCGCGCTGTCTCACTGCTCCCACTGCGCAACTGCTCTG	2176
621	AspAlaLeuGlnGlyLysSerThrSerAlaGlyArgProAlaAspArg	637
2177	GGGGCGCTGCAGACAGCTCGCGCGCGCGCTTCCGGGCGGCTCCAAAGAG	2226
637	gValGlnArgValThrGlnAlaLeuArgSerAlaLeuAspSerCys...	652


```

475 aleuLeuLHISsSerAlaaspGlyAlaGlyTyrGluArgLeuValGly 492
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1674 TGTGCTCCCTACTACAGCCGATGACGTGGGTTTCAGAGCCGCTGGGGGG 1723
492 lalaLeuAlaSerAlaLeuSerGlnMetProLeuArgValAlaValaspLeu 508
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1724 CCTGGGCTGGGCGCTGTGCTGACGCTGGCGGCGCTGGCGGCTGACCTG 1773
509 TPSPeArArGluLeuSerAlaHisGlyAlaLeuAlaTTPPhHisH 525
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1774 TGGAGCGGCTGACTGACGAGCCGAGGGGCGCTGGCTTGGTTACAGCC 1823
535 sGlnArGArGArgGlyLeuGlnGlnGlyValAlaValLeuLeuPhe 542
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1824 CGAGCGGCGCCAGACCTGCAGAGAGGCGGCTGGTGTCTTGGCTTCT 1873
542 ePrAlaAlaValAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 558
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1874 CTCGGGCTGGGCTGGCTGACGAGTGGCTGACAGATGGGGTGTCC 1923
559 GluPro.....GlyProHisAspAlaLeuAlaAlaTTPLeuSerCy 572
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1924 GGGCGCGGCGGAGGCGCGGAGAGGCTTCCGCGCTGCTGCTGACGCTG 1973
572 sValLeuProAspPheLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 589
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1974 CGTGTGCGCGGCTTGTGAGGCGGCGGCGGCGGCGGCTGACGAGG 2023
589 aTyPhePhePheLeuHisProAspSerValProSerProPheArg 605
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2024 CCGTGTGACAGGCTGCTGACCGGAGCGGAGCGGCTTGTGCTGCTG 2073
606 ValAlaProLeuPheSerLeuProThrGlnLeuProAlaPheLeuAspAl 622
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2074 ACCGGCGCGCTTGCACACTGCGCTCCCACTGCGCTGCTGCGGCG 2123
622 aleuGlnGlyGlyCysSerThrSerAlaGlyArgProAlaAspArgVal 639
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2124 CCTGAGAGCGCTGCGGCGGCGGCTTCCGCGGCGGCTCCACAGAGAGCG 2173
639 luArgValThrGlnAlaLeuArgSerAlaAspSerCys.....Thr 653
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2174 AGCAAGTGTCCGCGGCGCTTGCACGAGCGCTGAGTACTTCCATGCC 2223
654 SerSerSerGluAlaProGly 660
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2224 CCGGGGAGACNTCCGCGCGGGA 2244

```

seq.name: /SIS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAF93875

seq.documentation_block: AAF93875 standard; cDNA: 2499 BP.

AAF93875:

23-MAY-2001 (first entry)

Human cDNA encoding a membrane or secretory protein clone PSEC0233.

Human: secretory protein; membrane protein; vaccine; gene therapy; rheumatoid arthritis; diabetes; ss.

Homo sapiens.

EP1067182-A2.

10-JAN-2001.

07-JUL-2000; 2000EP-0114090.

08-JUL-1999; 99JP-0194179.

11-JAN-2000; 2000JP-0118775.

```

PR 02-MAY-2000; 2000JP-0183766.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
XX WPI: 2001-093989/11.
XX P-PSDB; AAB88448.
XX
XX Nucleic acids encoding secretory proteins/membrane proteins, useful in
PT gene therapy or as candidate target molecules in drug development -
XX
XX Claim 1: SEQ ID 263; 609pp + CD ROM; English.
XX
XX This invention relates to nucleic acid sequences AAF93744 - AAF93916
XX which encode human secretory or membrane proteins represented by
XX AAB88317 - AAB88419. Included in the invention are primers
XX AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
XX cDNA sequences of the invention. The invention also includes methods for
XX the production of antibodies directed against the proteins, and cDNA
XX sequences, which can be used in vaccines. The polynucleotide sequences
XX can be used in gene therapy. The polynucleotide sequences and the
XX proteins they encode may be used in the prevention, treatment and
XX diagnosis of diseases associated with inappropriate secretory
XX protein/membrane protein expression. The nucleic acids and complementary
XX sequences may also be used as DNA probes in diagnostic assays
XX (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
XX presence of similar nucleic acid sequences in samples. They may also be
XX used to study the expression and function of secretory proteins/membrane
XX polypeptides and their role in metabolism. The polypeptides may be used
XX as antigens in the production of antibodies against them and in assays to
XX identify modulators (agonists and antagonists) of expression and
XX activity. The antibodies and antagonists may also be used as therapeutic
XX agents to down regulate expression and activity. The antibodies may also
XX be used as diagnostic agents for detecting the presence of the
XX polypeptides in samples (e.g. by enzyme linked immunosorbent assay
XX (ELISA). Examples of diseases which may be treated include rheumatoid
XX arthritis and diabetes.
XX
XX Sequence 2499 BP; 418 A; 821 C; 764 G; 496 T; 0 other;

```

alignment_scores:

Quality: 2182.00

Length: 763

Ratio: 3.974

Gaps: 10

Percent Similarity: 71.953

Percent Identity: 58.060

alignment_block: US-09-899-471-2 x AAF93875

Align seg 1/1 to: AAF93875 from: 1 to: 2499

```

1 MetProValSerTriPheLeuLeuSerLeuAlaLeuGlyArgAsnProVal 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
208 ATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 257
17 lValValSerLeuGlnArgLeuMetGluProGlnAspThrAlaArgCys 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
258 GGTCTTCTCTGAGAGAGCTTGTGGGCTGTGAGAGAGAGAGAGAGAGAGAG 307
34 erLeuGlyLeuSerCysHisLeuTPAspGlyAspValAlaCysLeuPro 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
308 CTCCGGCGCTCTCTGAGAGAGCTTGTGGGCTGTGAGAGAGAGAGAGAGAG 357
51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuG 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
358 GGGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 407
67 nThrGluLeuValLeuArgCysProGlnLysThrAspCysAlaLeuArg 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
408 GACAGAGCTGTGCTGAGGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 457
84 alArgValValValHisLeuAlaValHisGlyHisThrAlaGluProGln 100

```



```

458 TGGGTGGCTGTCACCTGGCCGTGCATGGGCACTGGGAAGAGCTGA 507
101 Glu.....AlaGlyLysSerAspSerGluLeuGlnGlnSer 113
508 GATGAGAAAAGTTTGGAGAGCAGCTGACTCAGGGGTGAGAGACCTAG 557
113 GAsnAspSerLeuGlnAlaGlnValValLeuSerPheGlnAlaTyrProI 130
558 GAATGCTCTCTCCAGGCCCAAGTCGTGCTCTCTCCAGGCTCAACCTTA 607
130 LAlaArgCysAlaLeuLeuGlnValGlnValProAlaAspLeuValGln 146
608 CTGCGCGGTGCGTCTGCTGGAGTGCAGTGGCTGCTGCTTGTGCAG 657
147 ProGlyGlnSerValGlySerAlaValPheAspCysPheGlnAlaSerLe 163
658 TTTGGTCAGCTGTGGGCTGTGGTATGACTGCTTCGAGGGCTGCT 707
163 uGlyAlaGlnValGlnLeuPheSerTyrThrLysProArgTyrGlnLysG 180
708 AGGGAGTGAGTACGAATGTGCTTACTCTACGCCAGGTACAGAGAAG 757
180 LLeuAsnLeuThrGlnGlnLeuPro..... 188
758 AACTCAACACACACAGACGCTGCTGCTGCTGCTGCTCAACGTGCA 807
189 ...AspGlyAspAsnValLeuLeuThrLeuAspValSerGlnGlnAs 204
808 GCAGATGTGACACACGTGATCTGTTCTGAATGTCTCTGAGGAGCA 857
204 PheSerPheLeuLeuTyrLeuArgProValProAspAlaLeuLysSerL 221
858 CTTCGGCTCTCTGTTACTGGAATCAGGTCCAGGGCCCCCAAAACCCC 907
221 euTyrTyrLysAsn..... 225
908 GGTGGCACAAAACCTGTGAGGCTCTCCCTCCCAAGTCCATTCCAC 957
225 ..... 225
958 TGTAGCCGATGCTGTGCAAAAGACGAGTCCATATCAGAGATCC 1007
225 ..... 225
1008 TTGAAGAGGACTCACCCCAAGGAAAAATTGTGGGGAACTTCTGC 1057
225 ..... 225
1058 CTTCGTGTTCTTGACTTGGCCCTCCCTCTCTCTTATCTTCTC 1107
226 .....Leu.ThrGlyProGlnAsnIleThrL 234
1108 CAACCTCTCTCTTATTATTGTTCCACAGACTGAGCCGAGATCATTA 1157
234 euAsnHisThrAspLeuValProCysLeuCysIleGlnValTyrSerLeu 250
1158 TGAACACACAGACTAGTCCCTGCTCTGTATTCAAGGTGGCCCTG 1207
251 GluProAspSerGluArgValGluPheCysProPheArgGlnAspProG 267
1208 GAACCTGACTCGTTAGAGACAACTGTGCCCCCTTACAGGAGAACCCCG 1257
267 YAlaHisArgAsnLeuThrPheHisIleAlaArgLeuArgValLeuSerProG 284
1258 CGCACACCAAGAACTCTGCAAGCCGCGACTGCGACTGTGACCTGC 1307
284 LValTyrPheGlnLeuAspAlaProCysCysLeuProGlyLysValThrLeu 300
1308 AGACTGTGCTCTGACCGACCGCTGCTGCTGCTGCTGCTGCTGCTG 1357
301 CysTyrPheAlaProAspGlnSerProCysGlnProLeuValProProVa 317
|||||.....

```

```

1358 TGTGGCGGCTCCGGGTGGGAGACCCCTGCCAGCCACTGTCTCCACGCT 1407
317 LProGlnLysAsnAlaThrValAsnGluProGlnAspPheGlnLeuVal 334
1408 TTTCCGGAGAAACGTACGTGTGACAGAGTTCGAGTTCCATTGCTGA 1457
334 LArgHisProAsnLeuCysValGlnValSerThrTyrGlnLysValGln 350
1458 AAGCCACACCTTAACCTGTGTTCAGGTGACAGACGTCCGGAAAGCTG 1507
351 LeuGlnAlaCysSerTyrAlaAspSerLeuGlyProPheLysAspPme 367
1508 CTGCAGAGTGTCTTGGGCTGACTCCCTGGGGCTCTCAAAAGCATGT 1557
367 LLeuLeuValGluMetLysThrGlyLeuAsnAsnThrSerValCysAlaL 384
1558 GCTACTGTGGAGACACAGAGGCCCCAGAGACACAGATCTCTGTGCT 1607
384 euGluProSerGlyCysThrProLeuProSerMetAlaSerThrArgAla 400
1608 TGAACCCAGTGGGTGTACTTACTACCAACGCAAGACCTCCACG..... 1651
401 ALaArgLeuGlyGlnLeuLeuGlnAspPheArgSerHisGlnCysMe 417
1651 ..... 1651
417 LLeuThrAspAspAsnMetGlySerLeuThrAlaCysProMetA 434
1652 ...CTATGG...GACGATGACTTGGAGCGCTATGAGGCTGCGCCATGG 1694
434 sPlyTyrIleHisArgArgTyrValLeuValTyrPheLysLeuLeu 450
1695 ACAATATCATCCACAGAGCGGTGGGCCCTGCTGTGGCTGCTGCTACT 1744
451 LeuAlaAlaAlaLeuPhePhePheLeuLeuLysLysAspArgArgL 467
1745 TTTGGCTGCTGCTTCTCCCTCATCTCTTCAAAAAGATCAGCGAA 1794
467 s.....AlaAlaA 470
1795 AGGGTGGTGAAGCTTTGAACAGAGACGTCCGCGGGGGCGCGCA 1844
470 rGlySerArgThrAlaLeuLeuHisSerAlaAspGlyAlaGlyTyr 486
1845 GGGGC...CGCGGGCTGTGCTCTCTACTCAGCCGATGACTCGGGTTTC 1891
487 GluArgLeuValGlyAlaLeuAlaSerAlaLeuSerGlnMetProLeuArg 503
1892 GACCGCTGTGGCGCCCTGGCGCTGGCCCTGTGCCAGTCCGCTGGC 1941
503 GValAlaValAspLeuThrPheSerArgArgGluLeuSerLanHisGlyAlaL 520
1942 CTTGGCCCTTAACCTGTGAGCCGTCTGAACCTGAAGCCGCGGCGCGC 1991
520 euAlaThrPheHisAsnGlnArgArgIleLeuGlnGluGlyVal 536
1992 TGGCTTGGTTCAAGCGCAGCGCGCCAGACCTGCAAGAGGCGCGCTG 2041
537 ValIleLeuLeuPheSerProAlaAlaValAlaGlnGlnGlnTyrLe 553
2042 GTGGTCTCTCTTCTCTCCGTGGCGTGGCTGGCTGGCTGGCTGGCT 2091
553 uGlnLeuGlnThrValGluPro.....GlyProHisAspAlaLeu 567
2092 ACAAGATGGGGTTCGGGGCGCGGGGGCGCACGGCGCGACAGCGCTTCC 2141
567 LAlaTyrPheSerCysValLeuProAspPheLeuGlnGlyArgAlaThr 583
2142 GCGCTGCTGCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2191
584 GlyArgTyrValGlyValTyrPheAspGlyLeuLeuHisProAspSerVa 600
2192 GGCAGCTACGTGGGGCTGCTTGCACAGGCTGTCCACCGCGAGCGCT 2241

```


[illegible][illegible]


```

351 LeuGlnAlaIacysSerThrAlaAspSerLeuGlyProPheLeuAspAspMe 367
      ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
701 CTCAGAGAGAGCTTGTGGGCTGAGCTCCCTGGGGCTCTCAAGACATG 750
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
367 tLeuLeuValGlnMetLysThrGlyLeuAsnAspHisSerValCysAla 384
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
751 GCTACTTGTGGACACAGAGCCCCCAGACACAGATCCCTCTGTGCT 800
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
384 euGluProSerGlyCysThrProLeuProSerMetAlaSerThrAlaGala 400
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
801 TGGACCCAGAGGCTGACTACTACACCCAGCAAGCCCTCCACAGAGGCA 850
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
401 AlaArgLeuGlyGluGlnLeuGlnAspPheArgSerHisGlnCysMe 417
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
851 GCTGGCTTGGAGAGTACTTACTACAGACCTGACAGCCAGGTGCT 900
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
417 tGlnLeuTrpAsnAspAspMetGlySerLeuTrpAlaCysProMet 434
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
901 GCGAGCTATGG...GACGATGACTTGGAGCGCTATGGCTCCCTCCATG 947
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
434 sPlsYrYrIleHisArgArGTrpValLeuValTrpLeuAlaCysLeu 450
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
948 ACAATATACATCCACAGCGCTGGGCGCTGCTGGCTGCTGCCCTACTC 997
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
451 LeuAlaAlaAlaLeuPhePhePheLeuLeuLysLysAspArgArgLy 467
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
998 TTTGCGGCTGCGCTTCCCTCATCTCTCTCAAAAGATCACCGCA 1047
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
467 s.....AlaAla 470
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1048 AGGCTGCTAGGCTTGTAAACAGAGCTCCGCGGGCGCGCGCA 1097
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
470 rglYSerArgTrpAlaLeuLeuHisSerAlaAspGlyAlaGlyTr 486
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1098 GGGGC...CGCGGGCTGCTGCTCTCTACTCAGCGGATACCTCGGTTTC 1144
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
487 GluArgLeuValGlyAlaLeuAlaSerAlaLeuSerGlnMetProLeu 503
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1145 GACGCGCTGCTGGGCGCTGGCGCTGCGCTGCTGCTGCTGCTGCTG 1194
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
503 gValAlaValAspLeuTrpSerArgArgGluLeuSerAlaHisGlyAla 520
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1195 CGGCGCGCTGAGCTGAGACCGCTCGAAGCTGAGCGCGAGGGCGCG 1244
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
520 euAlaTrpPheHisHisGlnArgArgIleLeuGlnGlnGlyGlyVal 536
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1245 TGGCTTGGTTTCAAGCGAGCGCGCGACCTGTCAGAGGGCGCGT 1294
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
537 ValIleLeuLeuPheSerProAlaAlaValAlaGlnCysGlnGlnTr 553
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1295 GTGGCTTGTCTCTCTCTCCGCTGCGCTGCGCTGCGAGGAGTGGCT 1344
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
553 uGlnLeuGlnThrValGlnPro.....GlyProHisAspAlaLeu 567
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1345 ACAGGATGGGCTGTCCGGCGCGCGCGCGAGCGCGCGAGCTTCC 1394
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
567 lAlaTrpLeuSerCysValLeuProAspPheLeuGlnGlyArgAlaTr 583
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1395 GCGCTGCTGCTGAGCTGCTGCGCGCTTCTTGCAGGGCGGGCGCG 1444
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
584 GluArgTrpValGlyValTrpPheAspGlyLeuHisProAspSerVa 600
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1445 GGGAGCTACGTGGGCGCTGCTGACAGGCTGCTCCACCGCGAGCGCT 1494
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
600 lProSerProPheArgValAlaProLeuPheSerLeuProHisGln 617
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1495 ACCCGGCTTTCGCGACCGCGCGCTTACACTGCCCTCCCACTGC 1544
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
617 roAlaPheLeuAspAlaLeuGlnGlyCysSerThrSerAlaGlyArg 633
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1545 CAGACTTCTCGGGGCGCTGACAGACCTCGCGCGCGCTTCCGGGCGG 1594
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

634 ProAlaAspArgValGluArgValIleThrGlnAlaLeuArgSerAlaLeu
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1595 CTCACAGAGAGAGCGAGGCAAGTGTCCCGGCGCTTACAGCAGCCCTGG
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
650 pSerCys.....ThrSerSerSerGluAlaProGly 660
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1645 TAGCTACTTCCATCCCCCGGGGACTCTCCGCGCGGGA 1681
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NM2000.DAT:AAZ65269

```

```

seq_documentation_block:

```

```

ID AAZ65269 standard; DNA; 1752 BP.

```

```

XX AAZ65269;

```

```

DT 23-MAR-2000 (first entry)

```

```

DE Human secreted protein gene 20.

```

```

KW Human; secreted protein; cancer; tumour; developmental abnormality;

```

```

KW foetal deficiency; blood disorder; immune system disorder; inflammation;

```

```

KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;

```

```

KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;

```

```

KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;

```

```

KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;

```

```

KW therapy; chromosome 3; ds.

```

```

OS Homo sapiens.

```

```

PN W09958660-A1.

```

```

PD 18-NOV-1999.

```

```

PF 06-MAY-1999; 9900-US09847.

```

```

PR 12-MAY-1998; 9805-0085093.

```

```

PR 12-MAY-1998; 9805-0085094.

```

```

PR 12-MAY-1998; 9805-0085105.

```

```

PR 12-MAY-1998; 9805-0085180.

```

```

PR 18-MAY-1998; 9805-0085906.

```

```

PR 18-MAY-1998; 9805-0085920.

```

```

PR 18-MAY-1998; 9805-0085921.

```

```

PR 18-MAY-1998; 9805-0085922.

```

```

PR 18-MAY-1998; 9805-0085923.

```

```

PR 18-MAY-1998; 9805-0085924.

```

```

PR 18-MAY-1998; 9805-0085925.

```

```

PR 18-MAY-1998; 9805-0085925.

```

```

PR 18-MAY-1998; 9805-0085927.

```

```

PA (HUMA-) HUMAN GENOME SCI INC.

```

```

PI Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;

```

```

PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;

```

```

PI Lafleur DW, Endress GA, Edner R;

```

```

DR WPI; 2000-062296/05.

```

```

DR P-PSDB; AAY76143.

```

```

PT New isolated human genes and the secreted polypeptides they encode,

```

```

PT useful for diagnosis and treatment of e.g. cancers, neurological

```

```

PT disorders, immune diseases, inflammation or blood disorders

```

```

PS Claim 1; Page 308; 475pp; English.

```

```

XX AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.

```

```

XX This sequence was found to be present on human chromosome 3.

```

```

XX AAY76124 to AAY76223 represent the secreted proteins encoded by the 97

```

```

XX human genes. The genes and their corresponding secreted polypeptides are

```

```

XX useful for preventing, treating or ameliorating medical conditions,

```

```

XX e.g. by protein or gene therapy. Also pathological conditions can be

```

```

XX diagnosed by determining the amount of the new polypeptides in a sample

```

```

XX or by determining the presence of mutations in the new genes. Specific

```

```

XX uses are described for each of the 97 genes, based on which tissues they

```


101 SNCCNGCNYTNMSNTGYMGNYTNTGGGAYWSNGAYATHYTNNTGYTNCN 150

338 AsnLeuCysValGlnValSerThrTrpGluLysValGlnLeuGlnAlaCys 354


```

1948 CANGARNGCNCNGARCAGTWMNMGNCNCTNCARCNCNGCNTGTAAWS 1997
651 rCy.....ThSeSerSergIuaIaProgly 660
          :         :::::  |||||
1998 NTATYTYCAVCNCCNCGNMACNCCNCGNCGN 2031

seq_name: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAS18131

seq_documentation_block:
ID      AAS18131 standard; cDNA; 2109 BP.
XX
AC      AAS18131;
XX
DT      26-MAR-2002 (first entry)
XX
DE      Human DCRS7 reverse translation generic cDNA.
KW      Human; DNAX cytokine receptor subunit 7; DCRS7; phosphate labelling; ss;
KM      gene therapy; protein therapy; immunological disorder.
XX
OS      Homo sapiens.
XX
PN      WO200190358-A2.
XX
PP      29-NOV-2001.
XX
PR      23-MAY-2001; 2001WO-US16767.
XX
PS      24-MAY-2000; 2000US-206862P.
PA      (SCHE ) SCHERING CORP.
PI      Gorman DM;
PX
DR      WPI: 2002-106198/14.
XX
XX      Isolated antigenic human or mouse DNAX receptor subunit-like
PT      polypeptide useful for detecting antibodies generated in response to
PT      presence of increased protein levels or immunological disorders -
PS
PS      Disclosure: Page 15-16; 148pp; English.
XX
XX      The invention relates to primate and rodent DNAX cytokine receptor
CC      subunit (DCRS) polypeptides and the polynucleotides encoding them. The
CC      receptors, or their portions may be useful as phosphate labelling enzymes
CC      to label general or specific substrates. The subunits may also be
CC      functional immunogens to elicit recognising antibodies, or antigens
CC      capable of binding antibodies. A combination, e.g., including a DCRS can
CC      be used as an immunogen for the production of antisera or antibodies
CC      capable of distinguishing between other cytokine receptor family members.
CC      A purified DCRS can also be used as a reagent to detect antibodies
CC      generated in response to the presence of elevated levels of expression,
CC      or immunological disorders which lead to antibody production to the
CC      endogenous receptor. This sequence represents human DCRS7 reverse
CC      translation generic cDNA.
XX
SO      Sequence 2109 BP; 234 A; 267 C; 415 G; 269 T; 524 other;

alignment_scores:
Quality: 1548.00 Length: 706
Ratio: 3.625 Gaps: 7
Percent Similarity: 60.482 Percent Identity: 44.193

alignment_block:
US-09-899-471-2 x AAS18131 ..
Align seg 1/1 to: AAS18131 from: 1 to: 2109

1 MetProValSerTrpPheLeuLeuSerLeuAlaLeuGlyArgAsnProva 17
|||||  |||||  ::  |||  ||
1 ARGCGNGTCNTGGTTYYTNTYTNMSNYTGNCVYTNGMGNNMNSNCARTG 50

```


[illegible]

seq_name: /sids1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH9805
seq_documentation_block:
ID AAH9805 standard; cDNA, 1210 BP.
XX
XX AAH9805;
AC
XX
DT 16-OCT-2001 (first entry)
XX
DE Human protein encoding cDNA sequence SEQ ID NO:640.
XX
XX Human; Cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; vitruide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiagregant; haemostatic; vulnery; antifur; osteopathic; eczema;
KW dermatological; antiallergic; antisthmatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; hematopoietic disorder; platelet disorder; asthma;
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder; ss.
XX
XX Homo sapiens.
OS
XX WO200153455-A2.
PN
XX 26-JUL-2001.
PD
XX
PF 22-DEC-2000; 2000MO-US35017.
PE
XX
PR 23-DEC-1999; 99US-0471275.
PR 21-JAN-2000; 2000MS-0488725.
PR 25-APR-2000; 2000MS-0552317.
XX
XX (HXSE-) HXSEQ INC.
PA
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI: 2001-457603/49.
XX P-PDB; AAM25864.
DR
XX
PT Isolated human polynucleotides encoding Polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
PS Claim 1: Page 659-660; 1217pp; English;
XX
XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAH25963. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; vitruide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnery;
CC antiviral; osteopathic; dermatological; antiallergic; antisthmatic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,

CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.

XX Sequence 1210 BP; 227 A; 377 C; 352 G; 254 T; 0 other;

Alignment_scores:

Quality: 843.50 Length: 330
 Ratio: 3.749 Gaps: 7
 Percent Similarity: 68.182 Percent Identity: 55.152

Alignment_block:

US-09-899-471-2 x AAH99805

Align seg 1/1 to: AAH99805 from: 1 to: 1210

```

1 MetProValSerTrpPheLeuLeuSerLeuAlaLeuGluValArgAsnProVal 17
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
199 ATGCGTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 248

17 lValValSerLeuGluArgLeuMetClnProGlnAspThrAlaArgCys 34
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
249 GCTCCTTCTCTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 298

34 erLeuGlyLeuSerCysHisLeuTrpAspClnAspValLeuGluLeuPro 50
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
299 CTCCGGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 348

51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuGlu 67
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
349 GGGGACATCGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 398

67 nThrGluLeuValLeuArgCysProGlnTrpAspCysAlaLeuArgVal 84
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
399 GACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 448

84 aLArgValValValHisLeuAlaValHisGlyHisTrpAlaGluProGlu 100
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
449 TGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 498

101 Glu.....AlaGlyLysSerAspSerGluLeuGlnGluSerArg 113
  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
499 GATGAGGAAAGTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 548

113 gAsnAlaSerLeuGlnAlaGlnValValLeuSerPheGlnAlaTyrProI 130
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
549 GAATGCCCTCTCTCCAGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598

130 leAlaArgCysAlaLeuLeuGluValGlnValProAlaAspLeuValGln 146
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
599 CTGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 648

147 ProGlyGlnSerValGlySerAlaValPheAspCysPheGlnAlaSerLe 163
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
649 TTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 698

163 uGlyAlaGluValGlnIleTrpSerTrpThrLysProArgTrpGlnLysG 180
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
699 AGGAGTAGAGTACGAATCTGCTCTATCTACGAGCCAGGTACGAGAGAGG 748

180 luleuAsnLeuThrGlnGlnLeuPro..... 188
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
749 AACTCAACACACACAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798

189 .....As 189
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
799 AACAGCATCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 848
  
```

```

189 pGlyAspAsnValLeuLeuThrLeuAspValSerClnGluGlnAspPheS 206
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
849 TGGTGAACAGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 898

206 erPheLeuLeuTrpLeuArgProValProAspAlaLeuLysSerLeuTrp 222
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
899 GCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948

223 TyrLysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAspLe 239
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
949 CACAAACCTGCTGAGGCTCTCCCTCCCAAGTC...CATTCCTACTG 995

239 uValPro.CysLeuCys.....lleglValTrpSerLeu 250
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
996 TAGCCGATGCTGCTGCAAGAGCAGCAGTGCATGTCATGAGAGATCTT 1045

251 GluProAspSerGluArgValGluPheCysProPheArgGluAspProGlu 267
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1046 GAAGAGGACTCA.....CCCCA 1062

267 yAlaHisArgAsnLeuTrpHisIleAlaArgLeuArgValLeuSerProG 284
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1063 AGCAGGAGAAATGCTG..... 1081

284 lYValTrpGlnLeuAspAlaProCysCysLeuProGly 296
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1082 .....GGAACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1102

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA102817
seq_documentation_block:
ID AA102817 standard; DNA; 16491 BP.
XX
AC AA102817;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 5505.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
  cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN W020015320-A2.
XX
PD 02-AUG-2001.
XX
FF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
  
```


PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246539.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249254.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
XX		
DR	WPI; 2001-465570/50.	
XX		
PT	Isolated nucleic acid molecule encoding a reproductive system antigen	
PT	is used in preventing, treating or ameliorating a medical condition	
XX		
PS	Disclosure; SEQ ID NO 5505; 1297bp + Sequence Listing; English.	
CC	The present invention provides the protein and coding sequences of a	
CC	number of human reproductive system related antigens. These can be used	
CC	in the prevention and treatment of reproductive system disorders,	
CC	including cancer. The present sequence is a genomic sequence encoding a	
CC	protein of the invention.	
SQ	Sequence 16491 BP; 3837 A; 4251 C; 4180 G; 4223 T; 0 other;	
alignment_scores:		
Quality:	692.50	Length: 470
Ratio:	2.663	Gaps: 14
Percent Similarity:	55.319	Percent Identity: 39.574
alignment_block:		
US-09-899-471-2 x AAL02817		
Align seg 1/1	to: AAL02817 from: 1 to: 16491	
298 ValThrLeuCysTrpGlnAlaProAspGlnSerProCysGlnProLeuVal 314		


```

15096 GTTGTGTTTGTGTTTGTGAGAGAGTCTCACCTGTGCGCCAGGCTG 15145
314 IProProValProGlnLysAsnAlaThrValAsnGluProGlnAspPheG 331
15146 GAGTGCACAGGCGGCGATCTCGCTCACTGCACACTCTGCTCCGAGATTC 15195
331 InLeuValAlaGlyHisProAsnLeuCysValGlnValSerThrTrpGlu 347
15196 AAGCGATT .....CTCTTGCTTCAGCCTCC.....GAG 15224
348 LysValGlnLeuGlnAlaCysSerThrAlaAspSerLeuGlyPropely 364
15225 TAGCGGGATTACAGGCTGT.....GCCACACGCGCGGCTAATT.. 15266
364 sAspAspMetLeuLeuValGlnMetLysThrGlyLeuAsnAsnThrSery 381
15267 ...TTTGTATTTTTGTAGTAGAG.....ACGGGTTTACCGTGTAGCC 15306
381 aLcysAlaLeuGluProSeryGlyCysThrProLeuProSerMetAlaSer 397
15307 AGGATGGCTTGTAT.....GTCCCTGACC 15329
398 ThrArgAlaAlaArgLeuGlyGlnGluLeuLeuGlnAspPheArgSer.. 413
15330 TCGTATCTGCGCGCGCTCGGCTCCGGAAGTGTGGGATTATATAAAGGC 15379
414 .....HisGlnCys...MetGlnLeuTrp.. 420
15380 ATGAGCCACACCGACCGCGCAATTTCATTCTGTAGTCTCAAGTTTTC 15429
421 .....AsnAspAspAsnMetGlySerLeuTrp 429
15430 TTTCTGTGAGAGGTAAGTTTCTACTCTTCTCTGCGGTTC..... 15473
430 AlaCysProMetAspLysTyrIleHisArgArgTrpValLeuValTrpLe 446
15474 .....CCAGACATCCACCAAGCGCTGGGCGCTCGTGGGCT 15508
446 uAlaCysLeuLeuLeuAlaAlaAlaLeuPhePheLeuLeuLeuLysL 463
15509 GGCTGCTACTCTTTGCGCGTGGCTTCCCTCATCTCTTCTCAAA 15558
463 ysAspArgArgLys..... 467
15559 AGGATCAGCGAAAGGTAGCGCTTCCGCGTCCCATTCCTCCGCGGA 15608
467 ..... 467
15609 GGACCAAGTGGCTGTGGAGTTCTCGAGGAAAGCGGCGCGAGCTTC 15658
467 ..... 467
15659 ACTGGTGCCTCCGCGCTCTCCCTTAACAGGGTGGCTGAGCTTTGAA 15708
467 ..... 467
15709 ACAGAGCTCCGCTCGGGGGGTAGTGGAGACAGCGTGGCGGAGGCG 15758
467 ..... 467
15759 CGCCCCGGGAGCGGAGCTGTGCCAGCTCACTTCCCTCCCATCT 15808
468 .....AlaAlaArgGlySerArgThrAlaLeuLeuLeuH 479
15809 GTTTTTCGCGAGCGCGCGGAGGCG...CGCGGCTCTGCTCTCT 15855
479 IsSerAlaAspGlyAlaGlyTyrGluArgLeuValGlyAlaLeuAlaSer 495
15856 ACTCAGCGGATGACTCGGTTTCGAGCGCGCTGGGGCGCGCTGCGTGC 15905
496 AlaLeuSerGlnMetProLeuArgValAlaValAspLeuTrpSerArgAr 512
15906 GCCCTGTGCGAGCTGCGGCTGCGGCGCTAGACCTGTGAGAGCGCTGC 15955
512 gGluLeuSerAlaHisGlyAlaLeuAlaTrpPheHisGlnArgArg 529
15956 TGAAGTGAAGGCGAGGAGCGCGCTGCGCTGTGTTTACAGCGGAGCGGCG 16005
529 rGlyLeuGlnGlnGlyGlyValValIleLeuLeuPheSerProAlaAla 545
16006 AGACCTGAGAGAGGAGGCGGTGTGTCTTCTCTCTCCGAGTGC 16055
546 ValAlaGlnCysGlnGlnTrpLeuGlnLeuGlnThrValGluPro..... 560
16056 GTGGCGCTGTGCAGCAGTGGTACAGAGATGGGTTGCCGGCGCGGCG 16105
561 ...GlyProHisAspAlaLeuAlaAlaTrpLeuSerCysValLeuPro 576
16106 GCACGGCGCGCAGAGAGCGCTTCCGCGCTCGCTCAGCTGCGCTGCCG 16155
576 sPpHeuGlnGlnGlyArgAlaThrGlyArgTyrValGlyValTyrPheAsp 592
16156 ACTTCTTGAGGAGCGCGCGCGCGCGGCACTACGTGGGGGCTGCTGTGAC 16205
593 GlyLeuLeuHisProAspSeryAlaProSerProPheArgValAlaProLe 609
16206 AGGCTGTCTCCACCGGAGCGCGTACCAGCGGCTTTCGCGACCGTGCCT 16255
609 uPheSerLeuProThrGlnLeuProAlaPheLeuAspAlaLeuGlnGlyG 626
16256 CTTCACACTGCGCTCCCAACTGCCAGACTTCTGGGGGCGCTCAGCAGC 16305
626 LysSerSerThrSerAlaGlyArgProAlaAspArgValGluArgValThr 642
16306 CTGGGCGCGCGCTTCCGGGCGGCTCCAAAGAGAGCGGAGAGTGTCC 16355
643 GlnAlaLeuArgSerAlaLeuAspSeryCys.....ThrSerSerSergl 657
16356 CGGGCGCTTACCGCAAGCGCTGATAGCTACTTCCATCCCGGGGAGCTCC 16405
657 uAlaProGly 660
16406 CGCGCGGGA 16415

```



```

51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuG 67
|||||
349 GGAAGCCTCAGCTCTGCCAGGCGCTGTGCTAGTGTCCACCCCGCTCA 398
67 nThrGluLeuValLeuArgCysProGlnLysThrAspCysAlaLeuArg 84
|||||
399 GACGAGCGCTGCTGTAGGTCTCCACAGAAACAGATTGGCCCTCTGTG 448
84 AlaArgValValAlaHisLeuAlaValHisGlyHisTrpAlaGluProGlu 100
|||||
449 TCCGTGTGTGTGCTCCACTTGGCCGTGCATGGGCACCTGGCAGACCCTG 498
101 GluAlaGlyLysSerAspSerGluLeuGlnLysSerArgAsnAlaSerLe 117
|||||
499 GAAGCTGGAAGCTGTATCAGAACTCCAGAGCTCTAGGAACGCTCTCT 548
117 uGlnAlaGlnValValLeuSerPheGlnAlaTyrrProIleAlaArgCysA 134
|||||
549 CCAGGCGCCAGGTGTGCTCTCTCCAGGCTTACCCCATGCGCCGCTGTG 598
134 AlaLeuLeuGluValAlaGlnValProAlaAspLeuValGlnProGlyGlnSer 150
|||||
599 CCTCTGTGAGGTCCAGTGTGCCGCTGACTGTGTGACGCTGTGTGCTGCTCC 648
151 ValGlySerAlaValAlaPheAspCysPheGlnAlaSerLeuGluAlaGluVal 167
|||||
649 GTGGGTCTGTGCGGTATTGACTGTTCGAGGCTAGTCTTGGGGCTGAGGT 698
167 GlnIleTrpSerTyrrThrLysProArgTyrrGlnLysGluLeuAsnLeu 184
|||||
699 ACAGATCTGTGCTCTACACAGAACCCAGGTACAGAAAGACTCAACCTCA 748
184 hArgGlnLeuPro..... 188
749 CACAGACAGCTGCTGACTGCAGGGGCTTGAAGTCCGGGACAGCATCCAG 798
189 ..... AspGlyAspAsnVal 193
799 AGCTGCTGGGCTCTGCCCTGCTCAATGTGTCTACAGATGTGTGCAATGT 848
193 IleuLeuThrIleuAspValSerGluGlnIleuAspPheSerPheLeuLeu 210
|||||
849 CCTTGTGACACAGTGAATGTCTGTGAGAGACAGACTTACCTTCTTACTGT 898
210 TyrLeuArgProValProAspAlaLeuLysSerLeuTyrrLysAsnLeu 226
899 ACCTGCTGTCCAGTCCCGGATGCTCTCAAAATCCTTGTGTACAAAACCTG 948
227 ThrGlyProGlnAsnIlePheIleuAsnHisThrAspLeuValProCysLe 243
|||||
949 ACTGAGCTCTCAGAACATTACTTTAAACACACAGACTGTGCTCCCTGCT 998
243 uCysIleGlnValTrpSerLeuGlnProAspSerGluArgValAlaGluPheC 260
999 CTGCAATTAGGTGTGTGCTAGAGCCACACTGTAGAGGGTGTGAATTC 1048
260 TyrProPheArgGluAspProGlyAlaHisArgAsnLeuTrpHisIleAla 276
1049 GCCCTTCCGGGAAGATCCCGGTGACACAGGAACCTTGGCACATGTAGCC 1098
277 ArgLeuArgValLeuSerProGlyValTrpGlnLeuAspAlaProCysC 293
|||||
1099 AGGCTGCGGAGTACTGCTCCAGAGGATATGGACACTAGATCGCCCTGCTG 1148
293 sLeuProGlyLysValThrLeuCysTrpGlnAlaProAspGlnSerProC 310
|||||
1149 TCTGCGGGGCAAGTAACACTGTGCTGGCAGCAGCAGCAGACAGATCCCT 1198
310 yGlnIleProLeuValProProValProGlnLysAsnAlaThrValAsnGlu 326
|||||
1199 GCCAGCCACTTGTGTCCACAGTGTGCCACAGAAACCCACTGTGAATGAG 1248
327 ProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuCysValAlaGlnVal 343
|||||
1249 CCAGAGATTTCAGTTGGTGGCAGGCCACCCCAACCTCTGTGTCCAGT 1298
343 IserThrTrpGluLysValGlnLeuGlnAlaCysSerTrpAlaAspSerL 360
|||||
1299 GAGCACTGTGGAGAAAGTTTACGTCAAGCGTCTTGTGGGTGACTGCT 1348
360 euGlyProPheLysAspAspMetLeuLeuValGluMetLysThrGlyLeu 376
|||||
1349 TGGGGCGCTTCAAGATGTATATCTGTAGTGCAGATGAAGAAACCGGCTC 1398
377 AsnAsnThrSerValCysAlaLeuGlnProSerGlyCysThrProLeuPr 393
|||||
1399 AACACACATCAGTGTGTCTTGAACCAAGTGGCTGTACACCACTGCC 1448
393 oSerMetAlaSerThrArgAlaAlaArgLeuGlnGlyGlnLeuGlnAla 410
|||||
1449 CAGCATGGCTCCACAGAGCTGTCTGCGGAGAGAGATTCTGTCAAG 1498
410 sPheArgSerHisGlnCysMetGlnLeuTrpAsnAspAspAspMetGly 426
|||||
1499 ACTTCGATCAGACACCGATATACACTGTGGACATGACCAACATGGGA 1548
427 SerLeuTrpAlaCysProMetAspLysTyrrIleHisArgArgTrpValLe 443
|||||
1549 TCGCTATGGGCTGTGCCCATGAGCAAGTACATCCACAGGCGGTGGTCT 1598
443 uValTrpLeuAlaCysLeuLeuLeuAlaAlaAlaLeuPhePheLeuL 460
|||||
1599 AGTATGGCTGTGCTGTCTACTGTGCTGTGGCGCTTTCTTCTCTCC 1648
460 euLeuLysLysAspArgLysAlaAlaAlaArgLysSerArgTrpAlaLeu 476
|||||
1649 TTCTTAAAAAGACCCGACAGAAAGCCGCCGTGTCTCCGACAGGCTTG 1698
477 LeuLeuHisSerAlaAspGlyAlaGlyTyrrGluArgLeuValAlaLe 493
|||||
1699 CTCTCTCACTCCGCGACGAGCGGCTGAGACCGCTGTGGTGGAGACT 1748
493 uAlaSerAlaLeuSerGlnMetProLeuArgValAlaValAspLeuTrpS 510
|||||
1749 GCGTGTGCGGTTGAGCCAGATGCCACTGTGCGCTGCGCTGTGACTGTGGA 1798
510 eArgArgGluLeuSerAlaHisGlyAlaLeuAlaTrpPheHisHisGln 526
|||||
1799 GCGGCGCGAGGTGAGCGCGGACGAGGCCCTAGCCGTGTCCACACACAG 1848
527 ArgArgArgIleLeuGlnGluGlyValValIleLeuLeuPheSerPr 543
|||||
1849 CGACGCCGTATCTGTCCAGAGAGGTGCGGTATATCTTCTTCTGCGC 1898
543 GAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnLeuGlnThrValGluP 560
|||||
1899 CCGGCGCGTGGCGCAAGTGTACAGTGTGCTGCGCTGCGCTGCGACAGTGAAGC 1948
560 rGlyProHisAspAlaLeuAlaAlaTrpLeuSerCysValLeuProAsp 576
|||||
1949 CCGGCGCGATACGCCCTCGCGCGCTGCTCAGCTGCTGTGCTGATCCAGAT 1998
577 PheLeuGlnGlyArgAlaThrGlyArgTyrrValGlyValTyrrPheAspG 593
|||||
1999 TTCTGTCAAGGCGGGGCGACCGCGCTTACGTGCGGGTCTACTTCCAGCG 2048
593 yLeuLeuHisProAspSerValProSerProPheArgValAlaProLeuP 610
|||||
2049 GTTGTGTCAACCACTGTGTGCTTCCCGCTTCCCGCTGCGCGCTCT 2098
610 hSerLeuProThrGlnLeuProAlaPheLeuAspAlaLeuGlnGlyGly 626
|||||
2099 TCTCCCTGCGCTCGCAGCTGCGCGGCTTCTGTGATGCACATGCGAGGAGGC 2148
627 CysSerThrSerAlaGlyArgProAlaAspArgValGluArgValThrG 643
|||||

```


2149 TGCCTCACTTCGGGGGGGACCCCGGACCGGCTGAGACAGTACCA 2198
 643 nAlaLeuArgSerAlaLeuAspSerCysThrSerSerSerGluAlaProG 660
 2199 GCGCTGCGCGCTCGCGCTGAGACAGTACTTCTAGCTCGGAGGCCAG 2248
 660 lYcYcYcGslGluTTPAspLeuGlyProCysThrThrLeuGlu 674
 2249 GCTGCTGCGAGGAGTGGACCTGCGACCTGCACTACCTAGAA 2292

seq_name: gb:BC004759

seq_documentation_block:
 LOCUS BC004759 2015 bp mRNA linear ROD:12-JUL-2001
 DEFINITION Mus musculus, clone MGC:6973 IMAGE:3154616, mRNA, complete cds.
 ACCESSION BC004759
 VERSION BC004759.1 GI:13435809
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2015)
 DIRECT SUBMISSION
 Submitted (21-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgaps-rt@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 DNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/
 Contact: villalobos@bcm.tmc.edu.
 Villalobos, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAC Plate: 6 Row: C Column: 3.

Location/Qualifiers
 1. 2015
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="MGC:6973 IMAGE:3154616"
 /tissue_type="mammary tumor. Brca1-/l; MMTV-Cre model. 10
 months old, gross tissue."
 /clone_lib="NCI-CGAP_Mam3"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 177. 1880
 /codon_start=1
 /product="Unknown (protein for MGC:6973)"
 /protein_id="AA04759.1"
 /db_xref="GI:13435810"
 /translation="MPVSMFLSLALGRNPVVSLEIEMPRODARSGISGHLMD
 DYLCRGLSASAPVLPVLPRIOTELVIRPOTDCAICRVVYHILAVHGAEBEA
 ASIGAEVQIWSITKPRYQELNLTQDLPDCELEVRDSIOSCVLPMILVSTDSGJNL
 LTLVSEEDDFSLYLRVPDALKSLMTGPNITLHTDLVLCICIQVNSLEP
 DSEBVEFCFREDPAGHRLMLHRLVLSGVWQDAPCLPGKVTLCWQAPSDPC
 OPLVPPQKNATVNEPODOLVAGHNLGVSTWKKVLOALACTLADSDGAPSDML
 LVEMKGLNLTGYCALPESGCTPLPSMASTRARIGELLQDFRSHOCMOQMDNDNG
 SLMACGMKTYIHRRLVPLACILLAALFFLLKLDKRTSPSPGSAACYPISCKAG
 RPAATSGSTGCTCTTCLPPRSASPRSSPCPSRCLSMWHCREADLPKDDPRTGMN
 E"

CDS

BASE COUNT 379 a 616 c 595 g 425 t
 ORIGIN
 alignment_scores:
 Quality: 2995.50 Length: 698
 Ratio: 5.210 Gaps: 2
 Percent similarity: 82.378 Percent identity: 82.235

alignment_block:
 US-09-899-471-2 x BC004759
 Align seg 1/1 to: BC004759 from: 1 to: 2015

1 MetProValSerThrPheLeuLeuSerLeuAlaLeuGlyArgAsnProVal 17
 177 ATGCTGTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 226
 17 lValValSerLeuGluArgLeuMetGluProGlnAspThrAlaArgCys 34
 227 GGTGCTCTCTGAGAGAGACTGATGAGCCTCAGACACTGCACGCTGCT 276
 34 erLeuGlyLeuSerCysHisLeuTPAspGlyAspValLeuCysLeuPro 50
 277 CTCTAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 326
 51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuG 67
 327 GGAGCTTCCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 376
 67 nThrGluLeuValLeuArgCysProGlnThrAspCysAlaLeuArgY 84
 377 GACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426
 84 AlaArgValValAlaHisLeuAlaValHisGlyHisThrAlaGluProG 100
 427 TCCGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGA 476
 101 GluAlaGlyLysSerAspSerLeuLeuGlnGlnSerArgAsnAlaSerLe 117
 477 GAGCTGGAAGTCTGATTCAGACCTCCAGAGGTCTAGAGAGCCCTCTCT 526
 117 uGlnAlaGlnValValLeuSerPheGlnAlaThrProIleAlaArgCysA 134
 527 CCAGGCCAGGTGTGCTCTCTCTCCAGGCTTACCCACGCGCGCTGTG 576
 134 lAlaLeuGluValGlnValProAlaAspLeuValGlnProGlyGlnSer 150
 577 CCCTGCTGAGGTCCAGAGTCCCGGCTGACCTGCTGCTGCTGCTGCTGCT 626
 151 ValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGluVal 167
 627 GTGGTCTCTGCGGTATTCAGCTTTCAGAGCTTCTGCGGGCTGAGGT 676
 167 lGlnIleThrSerThrThrLysProArgTyrGlnLysGluLeuAsnLeuT 184
 677 ACAGATGTGCTCTACACGAGCCAGGATACAGAAAGAGCTCAACTCA 726
 184 hrcGlnGlnPro..... 188
 727 CACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 776
 189 AspGlyAspAsnVal 193
 777 AGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 826
 193 lLeuLeuThrLeuAspValSerGlnGlnGlnAspPheSerPheLeuLeuT 210
 827 CTTCTGACACTGATGCTCTGAGAGCAGACACTTGTAGCTTCTTACTGT 876
 210 yrLeuAspProValProAspAlaLeuLysSerLeuTyrTyrLysAsnLeu 226
 877 ACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926


```

227 ThrGlyProGlnAsnIleThrLeuAsnHisThrAspLeuValProCysLeu 243
|||||
927 ACTGACCTCAGAACACTTACTTTAAACACACAGACCTGCTCCCTGCT 976
243 uCysIleGlnValTyrSerLeuGlnProAspSerGluArgValGlnPheC 260
|||||
977 CTGCATTCAGGTGTGGTGGCTAGAGCCAGACTGTGAGAGGTGGAAATTC 1026
260 ysrProPheArgGlnAspProGlyAlaHisArgAsnLeuTyrPheIleAla 276
|||||
1027 GCCCCTCCGGGAGAGATCCCGGTGCACACAGAACTCTGGCACAATGACC 1076
277 ArgLeuArgValLeuSerProGlyValTyrGlnLeuAspAlaProCysCys 293
|||||
1077 AGGCTCCGGGTACTGTCCCGAGGGTATGGCAGCTGATGCGCTGTGCTG 1126
293 sLeuProGlyLysValThrLeuCysTyrGlnAlaProAspGlnSerProC 310
|||||
1127 TCTGGCGGGCAAGGTAAACACTGTGTGGGAGGACACAGACAGAGTCCCT 1176
310 yGlnIleProLeuValProProValProGlnLysAsnAlaThrValAsnGlu 326
|||||
1177 GCCAGCCTTGTGGCCACAGTGCCTCCAGAAAGAACCCACTGTGAAATGAG 1226
327 ProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuCysValGlnVa 343
|||||
1227 CCACAGATTTCCAGTTGGTGGGAGGACACCCCAACCTGTGTCCAGGT 1276
343 lserThrTyrGlnLysValGlnLeuGlnAlaCysSerTyrAlaAspSerL 360
|||||
1277 GAGCACTCTGGAGAAAGTTCAGCTGCAGCGCTGTGTGGCTGTGACTCCT 1326
360 euGlyProPheLysAspAspMetLeuLeuValGlnMetLysThrGlyLeu 376
|||||
1327 TGGGGCCCTTCAAGGATGATGCTGTGATGGAGATGAAGAACCGGCTC 1376
377 AsnAsnThrSerValCysAlaLeuGlnProSerGlyCysThrProLeuPr 393
|||||
1377 AACACACATTCAGTCTGTGCTTGGAAACCGAGTGGTGTACACCACTGGC 1426
393 oSerMetAlaSerThrArgAlaAlaArgLeuGlyGlnGlnLeuLeuGlnAla 410
|||||
1427 CAGCAGGCGCTCCAGCAGAGCTGCTGGCTGGGAGAGGACTGTGTCAG 1476
410 srPheArgSerHisGlnCysMetGlnLeuTyrAsnAspAspAspMetGly 426
|||||
1477 ACTTCGATCACAACAGTGTATGCAGCTGTGAACGATGACACATGGGA 1526
427 SerLeuTyrAlaCysProMetAspLysTyrIleHisArgArgTyrValLe 443
|||||
1527 TGGCTATGGGCTGCCCCCATGGACAGATACATCCACAGCGCTGGGTCT 1576
443 uValTyrLeuAlaCysLeuLeuAlaAlaAlaLeuPhePheLeuLeu 460
|||||
1577 AGATAGGTGGCTGCTGCTACTTGTGGCTGGGCTTTCTTCTCTCTCC 1626
460 euLeuLysLysAspArgArgLysAlaAlaArgLysArgTyrAlaLeu 476
|||||
1627 TTTCTAAAAA..... 1636
477 LeuLeuHisSerAlaAspArgLysAlaGlyTyrGlnArgLeuValGlyAlaLe 493
|||||
1636 ..... 1636
493 uAlaSerAlaLeuSerGlnMetProLeuArgValAlaValAspLeuTyrps 510
|||||
1636 ..... 1636
510 euArgArgGlnLeuSerAlaHisGlyAlaLeuAlaTyrPheHisHisGln 526
|||||
1636 ..... 1636

```

```

527 ArgArgArgIleLeuGlnGlnGlyValValIleLeuLeuPheSerPr 543
1636 ..... 1636
543 oAlaAlaValAlaGlnCysGlnGlnTyrPheGlnLeuGlnTyrValGlnP 560
1636 ..... 1636
560 roGlyProHisAspAlaLeuAlaAlaTyrPheSerCysValLeuProAsp 576
1637 ..GGACCGGATGACCCCTCGCGGCTGCTCAGCTGCGCTGCTGCTGAT 1684
577 PheLeuGlnGlyArgAlaThrGlyArgTyrValGlyValTyrPheAspG 593
1685 TTCCTGCAAGCGCGGCGGACCGGCGCTACGTGCGGCTACTTGCACGG 1734
593 yLeuLeuHisProAspSerValProSerProPheArgValAlaProLeuP 610
1735 GCTGCTGACCCAGACTGTGTGCTCCCGCTCCCGCTCCCGCTCCCTCT 1784
610 hesLeuProThrGlnLeuProAlaPheLeuAspAlaLeuGlnGlyGly 626
1785 TCTCCCTGCTCGCAGCTGCGCGCTTCTCTGTGATGCTGACAGGAGGC 1834
627 CysSerThrSerAlaGlyArgProAlaAspArgValGlnArgValThrG 643
1835 TGTCTCACTTCGCGGCGGCGGACCGCGGACCGGCTGACAGATGACCCA 1884
643 nAlaLeuArgSerAlaLeuAspSerCysThrSerSerSerGlnAlaProC 660
1885 GCGCGTGGCGGTCCGCGCTGACACAGCTGTACTTCTACTCGGAAGCCCG 1934
660 lCysCysGlnGlnTyrPAspLeuGlyProCysThrThrLeuGln 674
1935 GCTGCTGCGAGGAATGGAGACTGGGACCTGCTGACTACACTAGAA 1978
seq_name: gb-pat:AX350978
seq_documentation_block:
LOCUS AX350978 2094 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 12 from Patent WO0190358.
ACCESSION AX350978
VERSION AX350978.1 GI:18616354
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (sites)
AUTHORS Gorman,D.M.
TITLE Minimalian receptor proteins; related reagents and methods
JOURNAL Patent: WO 0190358-A 12 29-NOV-2001;
SCHERING CORPORATION (US)
FEATURES
source
1..2094
location/Qualifiers
BASE COUNT 247 a 251 c 393 g 288 t 918 others
ORIGIN
alignment_scores:
Quality: 2637.00 Length: 698
Ratio: 4.938 Gaps: 1
Percent Similarity: 76.504 Percent Identity: 69.198
alignment_block:
US-09-899-471-2 x AX350978
Align seg 1/1 to: AX350978 from: 1 to: 2094
1 MetProValSerTyrPheLeuLeuSerLeuAlaLeuGlyArgAsnProVal 17
|||||
::: ||| ||| ||| ||| |||

```


593 yLeuLeuHisProAspSerValProSerProPhaArgValAlaProLeuP 610
 1851 MYNYNYTTCACCCAGAYWNGTNCNMNSCNTTYWNGTNGCNCNCNTMYT 1900
 610 hSeSerLeuProThrGlnLeuProAlaPheLeuAspAlaLeuGlnGly 626
 1901 TYWSTNTNCNMNSCARYTNCNCNTTYTYNGAYGCMYTCNARGNGCN 1950
 627 CySeSerThrSerAlaGlyArgProAlaAspArgValGluArgValThrG 643
 1951 TGTWSTNACNMNSCNGCNGCNCNCNTTYNGTNGTNGTNGTNGTNGTNG 2000
 643 nAlaLeuArgSerAlaLeuAspSerCysThrSerSerSerGluAlaProG 660
 2001 RGCNTYNTMGWMSNGCNTNGAYWSTNTYACWMSWMSNMGARCCNCNG 2050
 660 lYcScyScyGluGluTPAspLeuGlyProCysThrThrLeuGlu 674
 2051 GNTGYTGYGARGARTGGAYTNGCNCNTGYACNACNTNGAR 2094
 seq_name: gb_pat:AX073993
 seq_documentation_block:
 LOCUS AX073993 2255 bp DNA linear PAT 06-FEB-2001
 DEFINITION Sequence 1 from Patent WO0104304.
 ACCESSION AX073993
 VERSION AX073993.1 GI:12710225
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 2255)
 TITLE Human cytokine receptor
 JOURNAL Patent: WO 0104304-A 1 18-JAN-2001;
 ZymoGenetics, Inc. (US)
 FEATURES
 source
 1..2255
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 154..2232
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC28435.1"
 /db_xref="GI:12710225"
 /translation="MPVWFLLSLALGRSPVLLSERLVGPDATHCSPGLSCLTMDSD
 DILCLPDIIVAPGVLAPTHIOTELVLRCKETDCDCLRAVHLAVHGHMEPEDE
 EKFGAAGSVEEPRNASLQAOVLSFOAYPTARCTDLEVOYPAALVORGOSGVVY
 DCFEALGSEVRIMSTQPRYEKELNHQQLPALPMANYSAGDNVHLVANSSEOHF
 GLSLYMNVOGPPRFRHKNLTGPOITITLNMHTDLPCLCIQVPLEPDSVTRNICPFR
 EDPRAHQLMQAARLRLLTLQSLMDAPCSLPAEALCMRAGGDCOPLVPLSWEN
 VTVKVLKPLKHPNLCVYNSEKQLQDCLNADSGPKLDVLLLETROPDR
 SLCALEPGCTSLSPKASTRAARLEGLYLDLOSOCLOLMDODDGLMALACPMKXTH
 KRMALVNLACLFLAALSLILLKKHAKAARGRAALLVYADSGEPFRVYCAASA
 LCOLPIRYAVDLMSRRELSSAGSPVAMPHARQOTLOEGSVVLLSPGVALISELQ
 DGVSGPRGKHGHDAFRASLSTVLPDIDGRAGSTVGACFDLHLHDVAPALERTVAV
 FTLSQLEDFEALQDPRAPRNSGRIDERRAEVSRALDPLDSYFHPHGPARGGVGP
 GAGGAGGGT"
 CDS
 369 a 748 c 703 g 435 t
 ORIGIN
 Alignment_scores:
 Quality: 2352.50 Length: 678
 Ratio: 4.149 Gaps: 5
 Percent Similarity: 83.628 Percent Identity: 67.404
 Alignment block:
 US-09-899-471-2 x AX073993
 Align seg 1/1 to: AX073993 from: 1 to: 2255

1 MetProValSerThrPheLeuSerLeuAlaLeuGluArgAsnProVal 17
 154 ATGCCCTGTCCTGCTGCTTCTGCTGCTTGGCACTGGGCGGAGCCAGT 203
 17 lValValSerLeuGluArgLeuMetGluProGlnAspThrAlaArgCys 34
 204 GGTCTTCTCTGAGAGGCTTGGGGCTCAGAGAGCTACCCAGTCT 253
 34 erLeuGlyLeuSerCysHisLeuThrAspAlaValLeuCysLeuPro 50
 254 CTCGGGCGCTCTCGCTGCTGCTGAGCAGTACATCTGCTGCTGCT 303
 51 GlySerLeuGlnSerLysArgProGlyProValLeuValProThrArgLeu 67
 304 GGGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 353
 67 nThrGluLeuValLeuArgCysProGlnLysThrAspCysAlaLeuArg 84
 354 GACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 403
 84 aArgValValValHisLeuAlaValHisGlyHisThrAlaGluProGlu 100
 404 TGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 453
 101 Gln.....AlaGlySerSerAspSerGluLeuGlnGluSerAr 113
 454 GATGAGCAAAAGTTTGGAGGAGCAGCTGACTCAGGGGTGAGGAGCTAG 503
 113 gAsnAlaSerLeuGlnAlaGlnValValLeuSerPheGlnAlaArgPro 130
 504 GATGCTCTCTCCAGGCCCAATGCTGCTGCTCTTCCAGGCCCTACCTTA 553
 130 lValArgCysAlaLeuLeuGluValGlnValProAlaAspLeuValGln 146
 554 CTGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 603
 147 ProGlyGlnSerValGlySerAlaValPheAspCysPheGluAlaSer 163
 604 TTTGTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 653
 163 uGlyAlaGluValGlnIleThrSerThrThrLysProArgTyrGlnLys 180
 654 AGGAGTGAAGTACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 703
 180 lLeuAsnLeuThrGlnGlnLeuPro..... 188
 704 AACTCAACACACACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 753
 189 ...AspGlyAspAsnValLeuLeuThrLeuAspValSerGluGlnAs 204
 754 GCAGATGCTGACAACTGCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 803
 204 PheSerPheLeuLeuThrLeuArgProValProAspAlaLeuLysSer 221
 804 CTTCGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 853
 221 euThrTyrLysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThr 237
 854 GGTGGCACAATAAAGCTGACTGACGACGACAGTATCTTGAACACACA 903
 238 AspLeuValProCysLeuCysIleGlnValTyrSerLeuGlnThrProAsp 254
 904 GACCTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 953
 254 rGluArgValGluPheCysProPheArgGluAspProGlyAlaHisArg 271
 954 CCTTAGAGAGCAATCTGCTGCTTCCAGGAGAGACCCCGGACACACAGA 1003
 271 snLeuThrHisIleAlaArgLeuArgValLeuSerProGlyValTyrGln 287
 1004 ACCTCTGGCAAGCCCGCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTG 1053
 288 LeuAspAlaProCysCysLeuProGlyLysValThrLeuCysThrPheAla 304


```

1054 CTGGAGGACCCGCTGCTGCGCCGAGAGCGGACCTGCTGCGCGGCG 1103
1055 AATGAGGACCCGCTGCTGCGCCGAGAGCGGACCTGCTGCGCGGCG
304 AProaspGlnSerProCysGlnProLeuValProProValProGlnLysA 321
1104 TCCGGGTGGGAGCCCTGCGGACGACGAGTCCACCCCTTCTCGGAGGA 1153
321 snAlaThrValAsnGluProGlnAspPheGlnLeuValAlaGlyHisPro 337
1154 AGGTACCTGTGGACAGAGCTTCTGAGTCCCATTTGCTGAAGGCCACCT 1203
338 AsnLeuCysValGlnValSerThrTrpGluLysValGlnLeuGlnAlaC 354
1204 AACCTCTGTTCAGGTGAACAGCTGCGAGAGCTGCAAGCTGCGAGAGTG 1253
354 sSerTrpAlaAspSerLeuGlyProPheLysAspAspMetLeuLeuValG 371
1254 CTGTGGGCTGACCTCCCTGGGCGCTCTCAAAAGAGAGTCTACTGTTGG 1303
371 LuMetLysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluProSer 387
1304 AGACAGAGAGCCGCCAGACAGACAGATCCCTGTGCTTGGAGGCCAGAT 1353
388 GlyCysThrProLeuProSerMetAlaSerThrArgAlaAlaArgLeuG 404
1354 GCGTGTACTTCACTACCCAGCAAAAGCCTCCAGAGGCGACCTGCGCTTG 1403
404 YGluGluLeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpA 421
1404 AGAGTACTTACTACAGACCTGCGAGCGACGAGTGTCTGCTGCTGCTAGG 1452
421 snAspAspAsnMetGlySerLeuTrpAlaCysProMetAlaLysTrpLe 437
1453 ..GAGGATGACTTGGAGCGGCTATGGGCGCTGCGCCCATGGACAAATAC 1500
438 HisArgArgTrpValLeuValTrpLeuAlaCysLeuLeuLeuAlaAla 1500
1501 CACAGAGCGTGGCGCCCTGCTGCGGCGCTGCGCTGCTGCTGCTGCTGCG 1550
454 AleuPhePheLeuLeuLeuLysAspArgArgLysAlaAlaArg 471
1551 GCTTCCCTCATCTCTCTCTCAAAAGATCAGCGGAAAGCGGCGCCCA 1600
471 LysArgTrpThrAlaLeuLeuHisSerAlaAspGlyAlaGlyTrpGlu 487
1601 GGGGCGGCGGCGCTGCTGCTCTACTCAGCGCATGACTGCGGTTTCAG 1650
488 ArgLeuValGlyAlaLeuAlaSerAlaLeuSerGlnMetProLeuArg 504
1651 CGCCGCTGGGCGGCGCTGCTGCGGCGCTGCTGCGGCTGCGGCGCTG 1700
504 AlaValAspLeuTrpSerArgArgLysLeuSerAlaHisGlyAlaLeu 521
1701 GCGCTGAGCTGTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1750
521 LysTrpPheHisHisGlnArgArgArgLysLeuGlnLysGlyAlaVal 537
1751 CTGGTTTCAACCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1800
538 IleLeuLeuPheSerProAlaAlaValAlaGlnCysGlnLysTrpLeu 554
1801 GTCTGCTCTCTCTCTGCGGCTGCGGCGGCTGCGAGGAGTGGGCTGCA 1850
554 nLeuGlnThrValGluPro.....GlyProHisAspAlaLeuAla 568
1851 GGAGGGGTGTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTTCCG 1900
568 LysTrpLeuSerCysValLeuProAspPheLeuGlnLysArgAlaTrpG 584
1901 CCTGCTCAGGCGGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1950
585 ArgTrpValGlyValTrpPheAspGlyLeuLeuHisProAspSerValP 601

```

```

1951 AGCTAGTGGGGGCTGCTTGCAGACGGCTGCTCCACCCGAGCCGTTAC 2000
601 oSerProPheArgValAlaProLeuPheSerLeuProThrGlnLeuProA 618
2001 CGCCCTTTTCCGACCGTCCGCTTTCACACATGCTCCCACTGCGAG 2050
618 lAspLeuAspAlaLeuGlnGlyGlyCysSerThrSerAlaGlyArgPro 634
2051 ACTTCTGGGGGCGCTGCGACGCTGCGCGCCCGCGCTGCGGCGGCTC 2100
635 AlaAspArgValGluArgValThrGlnAlaLeuArgSerAlaLeuAsp 651
2101 CAAGAGAGAGCGGAGGAGAGTGTCCGCGGCGCTTACAGCAGCCGTA 2150
651 rCys.....ThrSerSerSerGluAlaProGly 660
2151 CTACTTTCATCCCGCGGAGCTCCGCGCGGGA 2184

seq_name: gb_pat:AX092430
seq_documentation_block:
LOCUS AX092430 2380 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 161 from Patent WO0116318.
ACCESSION AX092430
VERSION AX092430.1 GI:13444529
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2380)
Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A.,
Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Watanabe,C.K. and
Wood,W.I.
Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURN.
Genentech, Inc. (US)
FEATURES
Location/Qualifiers
source
1..2380
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT
411 a 776 c 743 g 450 t
ORIGIN

alignment_scores:
Quality: 2339.00 Length: 692
Ratio: 4.118 Gaps: 7
Percent Similarity: 82.081 Percent Identity: 66.329

alignment_block:
US-09-899-471-2 x AX092430 ...
Align seg 1/1 to: AX092430 from: 1 to: 2380
1 MetProValSerTrpPheLeuLeuSerLeuAlaLeuGlyArgAsnProVa 17
233 ATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 282
17 lValAlaSerLeuGluArgLeuMetGluProGlnAspPheHisAlaArgCys 34
283 GGTCTTCTTCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 332
34 erLeuGlyLeuSerCysHisLeuTrpAspGlyAspValLeuCysLeuPro 50
333 CTCCGGGCTTCTCTCCGCGCTTGGGACAGTACATCTGCTGCTGCTGCT 382
51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuG 67
383 GGGGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 432
67 nThrGluLeuValLeuArgCysProGlnLysThrAspCysAlaLeuArg 84

```



```

|||||.....:|||||
433 GACAGAGCTGGTGTAGGTGCACAGAGACCGACTGTGACTCTGTC 482
84 aAlaGValValValHisLeuAlaValHisGlyHisTrpAlaGluProGlu 100
:::|||||:|||||
483 TGGGTGTGGCTGCCACTGGCCGTGATGGCACTGGGAAGACCTGGAA 532
101 Glu.....AlaGlyLysSerAspSerGluLeuGlnGluSerAr 113
:::|||||:|||||
533 GATGAGGAAAAATTTCGAGAGACGTGACTCAGGGGTGAGAGACCTAG 582
113 GAsnAlaSerLeuGlnAlaGlnValValLeuSerPheGlnAlaTrpPro 130
|||||:|||||
583 GATGCTCTCTCCAGGCCCAAGTGTGCTCTCTCCAGGCTTACCTA 632
130 LeuAlaGlySAlaLeuLeuGlnValGlnValProAlaAspLeuValGln 146
|||||:|||||
633 CTGCCCCGCTGCTCTGCTGAGGTCAAGTCCCTGCTGCTGCTGCTG 682
147 ProGlyGlnSerValGlySerAlaValPheAspGlyPheGlnAlaSer 163
|||||:|||||
683 TTGTGCTAGCTGTGGCTGTGTGTGTATGACTGTTGAGGCTGCTCT 732
163 uGlyAlaGlnValGlnIleTrpSerTrpHisProArgTrpGlnLysG 180
|||||:|||||
733 AGGAGTGAAGTACGAATCTGTGCTATATCTACGCCAGCTACGAGAAG 782
160 LuLeuAsnLeuThrGlnGlnLeuPro..... 188
|||||:|||||
783 AACTCAACCCACACAGACGCTCCCTGCTGCTGCTCAAGCTGTCA 832
189 ...AspGlyAspAsnValLeuLeuThrLeuAspValSerGluGlnAs 204
|||||:|||||
833 GCAGATGGGACAACTGATCTGTGTGAATGCTCTGAGGAGGACGA 882
204 PheSerPheLeuLeuTrpLeuArgProValProAspAlaLeuLysSer 221
|||||:|||||
883 CTTCGGCTCTCCCTGCTGCTGATCAGGTCCAGGCCGCCCAAAACCC 932
221 eutTrpTrpLysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHis 237
|||||:|||||
933 GGTGGCACAAAAACCTGACTGGACCCAGATCACTTGAACACACACA 982
238 AspLeuValProCysLeuGlyIleGlnValTrpSerLeuGluProAsp 254
|||||:|||||
983 GACCTGTCTCCCTGCTGCTGTATCAGGTGTGCTCTGGAACCTGACTC 1032
254 rGluArgValGluPheCysProPheArgGluAspProGlyAlaHisArg 271
|||||:|||||
1033 CGTTAGGAGCAATCTGCCCTTCAGGAGGAGACCCCGCGCACACACA 1082
271 snLeuTrpHisIleAlaArgLeuArgValLeuSerProGlyValAla 287
|||||:|||||
1083 ACCTCTGGCAACCGCCGACTGCGACTGCTGACCTGCAGACTGGCTG 1132
288 LeuAspAlaProCysCysLeuProGlyLysValThrLeuCysTrpGln 304
|||||:|||||
1133 CTGGAGCGACCGTCTGCTGCTGCCGCAAGAGGCGCACTGTGTGGGG 1182
304 aProAspGlnSerProCysGlnProLeuValProValProGlnLys 321
|||||:|||||
1183 TCCGGGTGGGAGACCCCTGCCAGCACTGTCCACGCTTCTCTGGAGA 1232
321 snAlaThrValAsnGluProGlnAspPheGlnLeuValAlaGlyHis 337
|||||:|||||
1233 ACCTACTGTGACAAAGTTCGAGCTTCCATCTGCTGAAGGCCCACT 1282
338 AsnLeuCysValGlnValSerThrTrpGluLysValGlnLeuGlnLac 354
|||||:|||||
1283 AACCTCTGTGTTCAGTGAACAGCTCGGAAGAGCTCAGCTGCAAGAG 1332
354 sSerTrpAlaAspSerLeuGlyProPheLysAspAspMetLeuLeuVal 371
|||||:|||||
1333 CTGTGGGCTGACTCCCTGGGCGCTCTCAAAAGATGTGCTACTGTTG 1382
371 LuMetLysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluPro 387
|||||:|||||
1383 AGAGACAGAGGCCCCCAAGACACAGATCCCTGTGTGCTTGAACCACT 1432
388 GlyCysTrpProLeuProSerMetAlaSerThrArgAlaAlaArgLeu 404
|||||:|||||
1433 GGCTGACTTCACTACCCAGCAAAAGCTTCCACAGAGGAGGAGCTGTG 1482
404 yGluGluLeuLeuGlnAspPheArgSerHisGlnCysMetGluLeu 421
|||||:|||||
1483 AGAGTACTTACTACAGAGCTGCACTGACAGGAGGAGCTGTGCTAGT 1531
421 snAspAsnMetGlySerLeuTrpAlaCysProMetAspLysTrp 437
|||||:|||||
1532 ..GACATGACTTGGAGGCGCTATGGCCCTGCCATGAGCAAAATATC 1579
438 HisArgArgTrpValLeuValTrpLeuAlaCysLeuLeuAlaAla 454
|||||:|||||
1580 CACAAAGCTGGGCGCTGTGTGTGCTGCTGCTGCTGCTGCTGCTG 1629
454 aLeuPhePhePheLeuLeuLysLysAspArgArgLys..... 467
|||||:|||||
1630 GCTTCCCTCATCTCTCTCTCAAAAAGATCACGCCAAAGGCTGCTGA 1679
468 .....AlaAlaArgLysSerArg 473
|||||:|||||
1680 GCCTTTGAAACAGACGTCCTCGGGGGGCGCGCCGACAGGGGC... 1726
474 ThrAlaLeuLeuLeuHisSerAlaAspGlyAlaGlyTrpGluArgLeu 490
|||||:|||||
1727 GGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1776
490 lGlyAlaLeuAlaSerAlaLeuSerGlnMetProLeuArgValAlaVal 507
|||||:|||||
1777 GGGCGCTCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1826
507 sPLeuTrpSerArgArgGluLeuSerAlaHisGlyAlaLeuAlaTrpPhe 523
|||||:|||||
1827 ACCTGGAGAGCGTCTGTAAGTACGCGGAGGCGCGGCTGTGCTTT 1876
524 HisHisGlnArgArgArgLeuGlnGlnGlnGlnGlnGlnGlnGln 540
|||||:|||||
1877 CACGGCGACAGCGCGCACGCCCTGCAAGAGGCGGCTGTGCTGCT 1926
540 uPheSerProAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnLeu 557
|||||:|||||
1927 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1976
557 hValGluPro.....GlyProHisAspAlaLeuAlaAlaTrpLeu 570
|||||:|||||
1977 TGTCCGGGCGCGGCGGCGCACAGGCCGACAGACCTTCCGCGCTGCT 2026
571 SerCysValLeuProAspPheLeuGlnGlnArgAlaHisGlyArgTrp 587
|||||:|||||
2027 ACCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2076
587 lGlyValTrpPheAspGlyLeuLeuHisProAspSerValProSer 604
|||||:|||||
2077 GGGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2126
604 hArgValAlaProLeuPheSerLeuProThrGlnLeuProAlaPheLeu 620
|||||:|||||
2127 TCCGCAACGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2176
621 AspAlaLeuGlnGlyCysSerThrSerAlaGlyArgProAlaAsp 637
|||||:|||||
2177 GGGGCTGTGACAGCTGCGCGCGCGCTTCCGGGCGCTTCCAAAGAG 2226
637 gValGluArgValThrGlnAlaLeuArgSerAlaLeuAspSerCys... 652
|||||:|||||
2227 AGCGAGCAAGTGTCCCGGCGCTTCTGAGCAGCGCTGATGCTACTTCC 2276

```



```

1483 AGAGTACTACTACAAAGACCTGACATGAGCCAGTGTCTGACGATATG. 1531
421 snspaspanmetelgysertleutrpalaCysPrometAspLysTyrIle 437
1532 ..GAGCATGACTGGAGCGCTATGGGCTGCGCCATGATGACAAATACATC 1579
438 HisArgArgTrpValIleTrpLeuAlaCysLeuLeuLeuAlaAlaAla 454
1580 CACAGAGCGTGGGCGCTCGTGGTGGCTGGCTGCTGCTCTTGGCGCTGC 1629
454 alauphepheheleuleuleulysAspArgArgLys..... 467
1630 GCATTCCCTATCTCTCTCTCAAAAAGATCACCGGAAAGGTGGCTGA 1679
468 .....AlaAlaArgGlySerArg 473
1680 GGCTCTTGAACAGAGACGTCGCGTGGGCGCGCCGACAGGGGC...CGC 1726
474 ThrAlaLeuLeuHisSerAlaAspLysAlaGlyTyrGluArgLeuVa 490
1727 GCGGCTCTGCTCTCTCTACTACGCGATGATCGGGTTTCGACGCGCTGT 1776
490 lglAlaLeuAlaSerAlaLeuSerGlnMetProLeuArgValAlaVala 507
1777 GGGCGGCTGCGGCGTGGCGCTGTCGACGCTCCGCTGGCGCTGGCGCTAG 1826
507 spleutrpserArgArgGluLeuSerAlaHisGlyAlaLeuAlaTrpPhe 523
1827 ACCTGTGAGACCGCTGTGAACTGACGCGGAGGGCGCGTGGCTGTGTTT 1876
524 HisHisGlnArgArgArgIleLeuGlnGlnGlyValValIleLeuLeu 540
1877 CAGCGCGAGGGCGCCAGACCTGACAGAGGGCGCGGTGGTGTCTGCT 1926
540 upheserProAlaAlaValaIaGlnCysGlnGlnTrpLeuGlnLeuGlnT 557
1927 CTCTCTCCCGGCGGTGGCGCTGTGAGGAGTGGCTACAGATGGGG 1976
557 hvValGluPro.....GlyProHisAspAlaLeuAlaAlaTrpLeu 570
1977 TGCGCGGCGCGGGGCGCCAGCGCCGACGACGCTTCCGCGCTCGCTC 2026
571 SerCysValLeuProAspPheLeuGlnGlyArgAlaTrpGlyArgTyrVa 587
2027 AGCTGCGCTGCTCCCGACTTCTTCCAGGGCGCGCGCGCTACGCTAGCT 2076
587 lglValAlaTrpPheAspGlyLeuLeuHisProAspSerValProSerProp 604
2077 GGGGCGCTGCTGCGACAGGCTGCTCCACCGGAGCGCGCTACCGCTCTT 2126
604 heArgValAlaAlaProLeuPheSerLeuProThrGlnLeuProAlaPheLeu 620
2127 TCCGACACCGGCGCGCTTCCACACTGCTCCCAACTCCAGACTTCTCTG 2176
621 AspAlaLeuGlnGlyGlyCysSerThrSerAlaGlyArgProAlaAspArg 637
2177 GGGGCGCTGCGACGCTCGCGCGCGCTTCCGGCGCGCTCCCAAGAGAG 2226
637 gValGluArgValThrGlnAlaLeuArgSerAlaLeuAspSerCys.... 652
2227 AGCGGAGCAAGTGTCCCGGCGCTTCCAGCCAGCCCTGATAGCTACTTCC 2276
653 ..ThrsSerSerGluAlaProGly 660
2277 ATCCCCCGGGGACTCCCGCGCGGGA 2302

```

```

seq_name: gp_pr:BC006411
seq_documentation_block:
LOCUS BC006411 2507 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, clone MGC:10763 IMAGE:3606276, mRNA, complete cds.
ACCESSION BC006411
VERSION BC006411.1 GI:13623590

```

KEYWORDS MGC.
human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2507)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (09-Apr-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-rt@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzyzinski, Keta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRLM Plate: 13 Row: 1 Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction.

FEATURES
source
Location/Qualifiers

1..2507
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:10763 IMAGE:3606276"
/tissue_type="uterus, endometrium adenocarcinoma"
/clone_lib="NIH-MGC_44"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
198..1814
/codon_start=1
/product="unknown (protein for MGC:10763)"
/protein_id="AA06411.1"
/db_xref="GI:13623591"
/translation="MPVWFLLSLALGSPVVLSERLVPODATHCSPLSCLRLMDS
DILCLIPDIPAPGPAVLAPTHLOTETLVKCEKEDDLGLRVAVHLAVHGHEPEDE
EKFEGAGDSVEEPRNASLQAOVYFQVLPARACILAEVPAALVVGQSGSVYV
DCEFAAGSEVRIWYTOBPYREKELNQQLDALPLNYSAGDNYHVLVNSBEOR
GLSLVWYQGPAPRRMRHNLGTPTITLNTDLVPCLTQVWPELPSVRYNICEPR
EDPRHQNLMQARLRLTLTOSWLDAPCLDEALNLCWRARPGDCQPLVPLSMEN
TVVDVLEFPLKGNPNCVOYNSSEKLOLQCLNADSLGPKDVLILLETROPQNR
SLCALPEGCTSLSPKASRYARARGLGTYLQDLSGCGCLQIMDDDLGALVACPKDYIH
KRALVNLACILFAALSLILNLDKDHAGMRLRLKQDVRSGEEMQALGCGPPPSO
ACASSPLSPSVFSGSGRGPRGSAPLSR"

CDS
BASE COUNT
ORIGIN
464 a 804 c 772 g 467 t

alignment_scores:
Quality: 2316.50 Length: 724
Ratio: 4.071 Gaps: 8
Percent Similarity: 78.591 Percent Identity: 63.536

alignment_block:
US-09-899-471-2 x BC006411 ..
Align seg 1/1 to: BC006411 from: 1 to: 2507

223 TyrIysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAspLeu 239
 931 CACAAAACACCTGAGTGGACCCGACATCATTTACCTTGAACACACAGACCT 980
 239 ValProCysLeuGlySerIleGlnValTrpSerLeuGluProAspSerGluA 256
 981 GGTTCCTGCTCTGTATTCAGGTGGCTGCTTGAACCTGACTCCGTTA 1030
 256 rValGluPheCysProPheArgGluAspProGluYAlaHisArgAsnLeu 272
 1031 GGAGAACATGTCGCCCTTTCAGGGAGAACCCCGGACACCAAGAACCTTC 1080
 273 TrpHisIleAlaArgLeuArgValLeuSerProGluYAlaTrpGlnLeuAs 289
 1081 TGGCAACCCCGCCGACCTGACCTGACCTGACCTGACAGAGCTGGCTGTGA 1130
 289 PAlaProCysCysLeuProGlyLysValThrLeuGlyStrpGlnAlaProA 306
 1131 GGCACCTGCTGCTGCTGCCGCGAAGCGGACCTGTGGCGGCTCCGG 1180
 306 spcInSerProCysGlnProLeuValProProValProGlnLysAsnAla 322
 1181 GTGGGAGCCCTGCGACACCTGCTCCACCGCTTCTGGGAGAAATGTC 1230
 323 ThrValAsnGluProGlnAspPheGlnLeuValAlaGlnHisProAsnLe 339
 1231 ACTGTGGAC..... 1239
 339 uCysValGlnValSerThrTrpGluLysValGlnLeuGlnAlaCysSerT 356
 1240GTGAACAGCTCGAGAAAGCTGACCTGACGAGGAGCTTGT 1279
 356 rPAlaAspSerLeuGlyProPheLysAspAspMetLeuLeuValGluMet 372
 1280 GGGCTGACTCCCTGGGCTCTCAAGACGATGTCTACTGTGTGAGACA 1329
 373 LysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluProSerGlyCys 389
 1330 CGAGGCCCGCCGACACAGATCCCTGTGCTGCTGGAACCCAGTGGCTG 1379
 389 sThrProLeuProSerMetAlaSerThrArgAlaAlaArgLeuGlyGluG 406
 1380 TACTTCACATCCACCAAGAGCTCCACAGAGGAGCTGCCCTTGAGAGAT 1429
 406 LuLeuLeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsnAsp 422
 1430 ACTTACTACAGACCTCAGTCAAGGCCAGGTGTCTGCACTATGG...GAC 1476
 423 AspAsnMetGlySerLeuTrpAlaCysProMetAspLysTyrIleHisArg 439
 1477 GAGACATGTGGAGCGGTATGGGCTGCCCATGACAAATACATCATCAACA 1526
 439 gArgTrpValLeuValTrpLeuAlaCysLeuLeuLeuAlaAlaLeuP 456
 1527 GCGCTGAGGCTGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1576
 456 hepPheLeuLeuLeuLysLysAspArgLys..... 467
 1577 CCTCATCTCTCTTCTCAAAAAGATCAAGGAGGCTGAGGCTC 1626
 468AlaAlaArgGlySerArgThrAl 475
 1627 TTGAACAGACGTCGCTGGGGGCGGCGCCAGGCGC...CGCGCGC 1673
 475 AlaLeuLeuLeuHisSerAlaAspGlyAlaGlyTyrGluArgLeuValGlyA 492
 1674 TCTGCTCTCTTACTCAGCCGATGACTGGGTTTTCGAGCGCTGTGGGCGG 1723
 492 LeuAlaSerAlaLeuSerGlnMetProLeuArgValAlaAlaValAspLeu 508
 1724 CCTGGCGCTGGGCTGTGGCTGCGGCTGGCGCTGGCGCTGGAGACTGG 1773
 509 TrpSerArgArgGluLeuSerAlaHisGlyAlaLeuAlaTrpPheHisI 525

1774 TGAGCCGTCGTGAACTAGCGCGGAGGCGCCGCTGGCTTGTTCACCC 1823
 525 sGlnArgArgArgIleLeuGlnGlyGlyValValIleLeuLeuPheS 542
 1824 GCAGCGCGCGCCAGACCCCTGACAGAGGCGGCTGTGGCTGCTGCTTCT 1873
 542 erProAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnLeuGlnThrVal 558
 1874 CTCCGCTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1923
 559 GluPro.....GlyProHisAspAlaLeuAlaAlaTrpLeuSerCys 572
 1924 GCG 1973
 572 sValLeuProAspPheLeuGlnGlyArgAlaThrGlyArgValGlyV 589
 1974 CGTCTGCTCCGACATTTCTGACAGGCGCGCGCGCGCGCGCGCGCGCG 2023
 589 aLTrpPheAspGlyLeuLeuHisProAspSerValProSerProPheArg 605
 2024 CTGCTTTCGACAGGCTGCTCCACCGCGGACCGCTGACCCGCTTTTCCGC 2073
 606 ValAlaProLeuPheSerLeuProThrGlnLeuProAlaPheLeuAspAl 622
 2074 ACCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2123
 622 AlaGlnGlyGlyCysSerThrSerAlaGlyArgProAlaAspArgValG 639
 2124 CTGTCACAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2173
 639 LuArgValThrGlnAlaLeuArgSerAlaLeuAspSerCys.....Thr 653
 2174 AGCAAGTGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2223
 654 SerSerSerGluAlaProGly 660
 2224 CCGGCGACNTCCGCGCGCGGA 2244

seq_name: gb_pat:AX136341
 seq_documentation_block: 2499 bp DNA linear PAT 30-MAY-2001
 LOCUS AX136341
 DEFINITION Sequence 263 from Patent EP1067182.
 ACCESSION AX136341
 VERSION AX136341.1 GI:14272747
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2499)
 Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and
 Hayashi,K.
 TITLE Secretory protein or membrane protein
 JOURNAL Patent: EP 1067182-A 263 10-JAN-2001;
 Helix Research Institute (JP)
 FEATURES
 source
 1. 2499
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 208..1011
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC39812.1"
 /db_xref="GI:14272748"
 /translation="MPVRFELSLAIGRPVYLSERLVGPDDATCSPELSRLMS
 DILCLAGDIVPAGFVLAPTHIQTELVLVLCQKETDDILIRAVNHLAVHGMPEPDE
 EKFGAAGSGVEEPNNAIQADVLSFQAVPTPARCVLLEVOYFALALVFGSGSVVY
 DCFEALAGSEVAINSYTOIPRYEKELNHTQQLPALPLNLSADGDVNLVNLVSEEDHF
 GLSLYMNVOGPPKPRMKNLVRPPSOVHSHCRPMPVGRTOCHIREDP"

BASE COUNT 418 a 821 c 764 g 496 t
 ORIGIN

[illegible]

	seq_documentation_block:	2076 bp	DNA	linear	PAT 06-FEB-2001
LOCUS	AX073995				
DEFINITION	Sequence 3 from Patent WO0104304.				
ACCESSION	AX073995				
VERSION	AX073995.1 GI:12710227				
KEYWORDS					
SOURCE					
ORGANISM	synthetic construct. artificial sequence.				
REFERENCE	1 (bases 1 to 2076). Plesnell,S.R., Burkhead,S.K. and Pownder,S.L. Human cytokine receptor Patent: WO 0104304-A 3 18-JAN-2001; ZymoGenetics, Inc. (US)				
JOURNAL					
FEATURES					
source	location/qualifiers 1..2076 /organism="synthetic construct" /db_xref="taxon:32630" /note="This degenerate sequence encodes the amino acid sequence of SEQ ID NO:2."				
misc_feature	1..2076 /note="n = A,T,C or G"				
variation	1..2076				
BASE COUNT	232 a 273 c 399 g 265 t 907 others				
ORIGIN					
alignment_scores:					
Quality:	1652.50	length:	678		
Ratio:	3.764	Gaps:	5		
Percent Similarity:	64.749	Percent Identity:	47.640		
alignment_block:					
US-09-899-471-2 x AX073995					
Align seg 1/1	to: AX073995 from: 1 to: 2076				
1 MctProValSerTriPheLeuLeuSerLeuAlaLeuGlyArgAspProVa 17		::::			
1 ATGCCGGTNCNTGTGTTTYNTTNMNSNTNGCNYNMGNMWMSKCGT 50		::::			
17 lValAlSerLeuGluArgLeuMetGluProGlnAspThrAlaArgCys 34		::::			
51 NcTNTNMSNTNGARMGNTYNTGNCCNCARAGTCMACNCAYTGW 100		::::			
34 erLeuGlyLeuSerCysHisLeuTrpAspGlyAspValLeuCysLeuPro 50	::::				
101 SNCKNGCNTNMSNTNGMGNTYNTGGGAATWSNGAVATHYTNTGYTNCN 150	::::				


```

406 luLeuLeuGlnaspheargSerHisGlnCysmetGlnLeuTrpAsnasp 422
      ||||| ..... ||||| ||| |||
1250 AAYTYNTNCARAGAYTNCARMSNGCARGYTCARNTYNTGG...GAY 1296
      ||||| ..... ||||| ||| |||
423 AspAsnMetGlySerLeuTrpAlaCysPrometAspLysTrpIleHisar 439
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1297 GAYGAYTYTNGMGCGNTTNGGCGNTGCCATGGAARAAATVATHCAYAA 1346
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
439 gArGTrpValLeuValTrpLeuAlaCysLeuLeuLeuAlaAlaLeu 456
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1347 RMGNTGGCGGNTYNTGTYTNGCNGTYTNTYNTYNTYNTGCGCGCNGCNTN 1396
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
456 hepheheLeuLeuLeuLysAspArgArgLys..... 467
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1397 SMYTNATHYNTYNTNAAAGAYCAVCANAAAGCNGTGTNTMGNTN 1446
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
468 .....AlaAlaArgGlySerArgThrAla 476
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1447 YTNAAARCAAGATMGNGMNGCNGCNGCNGMNGCNGCNGCNGNT 1496
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
476 uLeuLeuHisSerAlaAspGlyAlaGlyTrpGluArgLeuValGlyAla 493
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1497 NYTNTNTATYWSNGCGATGAYWSNGGNTTGGARMGNTTGTNGCNGCNT 1546
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
493 euAlaSerAlaLeuSerGlnMetProLeuArgValAlaValaAspLeuTrp 509
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1547 TNGCMMNSGCGNTNTGTCARNTMCCNTYTNMGNGCNGTNGAYTYTWTGG 1596
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
510 SerArgArgGluLeuSerAlaHisGlyAlaAlaLeuAlaTrpPheHisHisG 526
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1597 WSNMGNGGATYTWMSNGCARGGCGCNGCNGTGTGTTCAYACGCGCA 1646
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
526 narGArgArgGlyLeuGlnGlyValValaIleLeuLeuPheSer 543
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1647 RMGNGNCARACNTNRCARGARGNGCNGTNGTNTYNTYNTYNTYNTWSNC 1696
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
543 roAlaAlaValaIaGlnCysGlnGlnTrpLeuGlnLeuGlnTrpValG 559
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1697 CNGGCGCNGCNGTNTNTGYSNGARTGTGTCARCAVGSNGTMSNGCN 1746
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
560 Pro.....GlyProHisAspAlaLeuAlaAlaTrpLeuSerCysVa 573
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1747 CGCGNGCNCAGCGCNCAGAYGCGTNTYMGCGCMMSTNTMSNTGTGT 1796
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
573 lLeuProAspPheLeuGlnGlyAlaGlnAlaThrArgTrpValGVal 590
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1797 NYTNCNGATYTYTNCARGMNGCNGCNGGMSNTATGTCNGCNGCNT 1846
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
590 yrPheAspGlyLeuLeuHisProAspSerValProSerProPheArgVal 606
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1847 GYTYGAYGNTYNTNCAVCCNGAYGNGTCNGCNGTNTNTYTMGNACN 1896
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
607 AlaProLeuPheSerLeuProThrGlnLeuProAlaPheLeuAspAla 623
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1897 GTTCCNGTNTTTCNTCCNMCNCAATYTCGATYTYTNGCNGCNT 1946
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
623 uGlnGlyGlyCysSerThrSerAlaGlyArgProAlaAspArgValGln 640
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1947 NCARARCCNMNGCNCNMNGMNSNGNGNTTNCARGARMGNGCNGARG 1996
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
640 rGValThrGlnAlaLeuArgSerAlaLeuAspSerCys.....ThrSer 654
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1997 ARGTMMSNMNGCNGTNCARCCNGCNGTNGATWSNTATYTCAYCCNCCN 2046
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
655 SerSerGluAlaProGly 660
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
2047 GGNACMNSNGCNCNGCN 2064
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||

```

seq_name: gb_pat:AX073998

seq_documentation_block:

```

12005 AX073998 1725 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 6 from Patent WO0104304.

```

```

ACCESSION AX073998
VERSION AX073998.1 GI:12710230
KEYWORDS
SOURCE
ORGANISM synthetic construct.
REFERENCE
AUTHORS Presnell,S.R., Burkhead,S.K. and Pownder,S.L.
TITLE Human cytokine receptor
JOURNAL Patent: WO 0104304-A 6 18-JAN-2001;
ZymoGenetics, Inc. (US)
FEATURES
source
location/Qualifiers
1..1725
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="This degenerate sequence encodes the amino acid
sequence of SEQ ID NO:5."
misc_feature
1..1725
/note="n = A,T,C or G"
variation
1..1725
BASE COUNT 192 a 226 c 332 g 216 t 759 others
ORIGIN
alignment_scores:
Quality: 1291.50 Length: 578
Ratio: 3.628 Gaps: 6
Percent Similarity: 61.592 Percent Identity: 45.502
alignment_block:
US-09-899-471-2 x AX073998
Align seg 1/1 to: AX073998 from: 1 to: 1725
110 GlnGluSerArgAsnAlaSerLeuGlnAlaGlnValValLeuSerPheG 126
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
1 GARGARCCMGNAAYCCMNSNTNRCARGCNCARGTGTNTNTMSNTTYCA 50
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
126 nAlaTrpProIleAlaArgCysAlaLeuLeuGlnValGlnValProAla 143
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
51 RCGNTAYCCNACNGCMGTGTNTYNTNRCARGTNCARGTNCNGCNG 100
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
143 splLeuValGlnProGlyGlnSerValGlySerAlaValPheAspCysPhe 159
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
101 CNTYNTNCARTTYGNCARMSNGTNGMNSGNTNTNATYAGATGTTY 150
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
160 GluAlaSerLeuGlyAlaGlnValGlnIleTrpSerTrpThrLysProAr 176
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
151 GARGCNGCCTYTNMGMSNARGTNGMGNATHGTGMSNTATACNARCCNM 200
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
176 gTyGlnLysGluLeuAsnLeuThrGlnLeuPro..... 188
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
201 NTAYGARAARGARYTAAAYCAVANCARCARLYTNCNGCYTNCNTG 250
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
189 .....AspGlyAspAsnValLeuLeuThrLeuAspValSer 200
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
251 TNAAYGTMSNGCNGAYGNGAYAAAGTNCAYTYTNTNTNAAVGTMSN 300
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
201 GluGlnGlnAspPheSerPheLeuLeuTrpLeuArgProValProAspAl 217
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
301 GARGARCARCATYTYGNTTMSNTNTNTATYTGAAATCARGTNCARGCNC 350
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
217 aLeuLysSerLeuTrpTrpLysAsnLeuThrGlyProGlnAsnIleThrL 234
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
351 NCCNARCCNMNGTNGCAVAAARAAYTTCNCGCNCNCARATATACN 400
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
234 euAsnHisThrAspLeuValProCysLeuCysIleGlnValTrpSerLeu 250
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
401 TNAACAYACNCAAYTYTNGTNCNTGYTNTGYATHTCARGTNTGCCNTY 450
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||
251 GluProAspSerGluArgValGluPheCysProPheArgGluAspProGl 267
      ||||| ..... ||||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

451 GARCNGATYMSNGTNGMNCNAATHTGYCCTWTYMGNGARAGAYCCNM 500
267 YLALHSHASGASNLLeuTriPhisILeAlaArgLeuAlaGValLeuSerPro 284
501 NCNCNGATYMSNGTNGMNCNAATHTGYCCTWTYMGNGARAGAYCCNM 550
284 YLALHSHASGASNLLeuTriPhisILeAlaArgLeuAlaGValLeuSerPro 300
551 ARMSNTGTTNTNGATGNCCTGMSNTTNCNGCNGARAGAYCCNM 600
301 CysTTPGAlaAlaProAspGlnSerProCysGlnProLeuValProProva 317
601 TGYTGNGMNCNGCNGCNGCNGCNGCNGCNGCNGCNGCNGCNGCNGCNG 650
317 ProGlnLysAsnAlaThrValAsnGluProGlnAspPheGlnLeuVala 334
651 MMSNTGGARAAAGTACNGCTNGAY..... 675
334 LAGLYHISProAsnLeuGlyValGlnAlaSerThrTrpGluLysValGln 350
676 .....GTNAATYMSNGARAAATYTCAR 699
351 LeuGlnAlaCysSerTrpAlaAspSerLeuGlyProPheLysAspAspMe 367
700 YTCARAGARTGYTNTGGCNGAYMSYNTGNCNTNARAGAYGVT 749
367 LeuLeuValGluMetLysThrGlyLeuAsnAsnTrpSerValCysAlaL 384
750 NYTNTNTNGARACNMGNCNCARAGAYAAAYMGMSNTYNTGYCNY 799
384 euGluProSerGlyCysThrProLeuProSerMetAlaSerThrArgAla 400
800 TNGARCCNMNSNGNTGACNWSNTYTCNMNSNARCCNMNSNACNMNGCN 849
401 AlaArgLeuGlyGluLeuLeuGlnAspPheArgSerHisGlnCysMe 417
850 CGNMGYTNGMNCARATAYTNTNARAGAYTCARMSNGCNCARTGYT 899
417 tGlnLeuTrpAsnAspAspMetGlySerLeuTrpAlaCysTrpMet 434
900 NCARTYNTG...GAYGAYGAYTNGCNGCNYTNGGCTGTCNATG 946
434 sPLysTyrILeHisArgTrpValLeuValTrpLeuAlaCysLeuLeu 450
947 ATRARATVTHCAIARMGNTGGCNYTNTGTYNGCTGTYTNTN 996
451 LeuAlaAlaAlaLeuPhePheLeuLeuLysLysAspArgArgLys 467
997 TTYGCGCNGCNYTNTNATHTYNTYTNARARAGAYCCNAA 1046
467 s.....AlaAlaArg 471
1047 RCGNTGGTYNMGNNTYTNARARAGAYTNGMNSNGCNGCNGCNM 1096
471 LysSerArgThrAlaLeuLeuLeuHisSerAlaAspGlyAlaGlyTyrGlu 487
1097 GNGGNGNCNCNTNTNTYTNATYMSNCNCAYGAYMSNGCTTGAR 1146
488 ArgLeuValGlyAlaLeuAlaSerAlaLeuSerGlnMetProLeuArgVa 504
1147 MGNNTGTYNGMNCNTNGCNGCNGCNYTNTGTYTNTGCTGTYTNT 1196
504 LAlaValaAspLeuTrpSerArgArgGluLeuSerAlaHisGlyAlaLeu 521
1197 NCNCNGNGAYTNTGMSNMGNMGNARCTYTNMNGCNGCNGCNGCTG 1246
521 LsTrpPheHisHisGlnArgArgGlyLeuGlnGluGlyValVal 537
1247 CMTGCTTAYACNCARMGNCARACACNTTNCARAGGNGGNGTGTN 1296
538 LLeuLeuPheSerProAlaAlaValaGlnCysGlnGlnTrpLeuGlu 554
1297 GNTYNTYNTYMSNCNGCNGCNGCNGCNYTNTGYMSNARAGGTYNCA 1346

```

```

554 LLeuGlnTrpValGluPro.....GlyProHisAspAlaLeuAla 568
1347 RGATGNGTNGMNSNGCNGCNGCNGCNGCNGCNGCNGCNGCNGCNG 1396
568 LsTrpLeuSerCysValaAlaProAspPheLeuGlnGlyArgAlaThrGly 584
1397 CMTSNTTMSNTGYTNTNCCNGATYTTTNCARAGGNGMNGCNGCNGN 1446
585 ArgTyrValGlyValaLysTrpPheAspGlyLeuLeuHisProAspSerValPr 601
1447 MNTATAGTNGMNCNTGYTNGAYMGNNTYTNATCAICCNATYGCNGNC 1496
601 oSerProPheArgValaAlaProLeuPheSerLeuProThrGlnLeuPro 618
1497 NCNTYNTTYMGNACNTGNCNTNTTYTACNTYTCNMNSNCARTYNCNG 1546
618 LsPheLeuAspAlaLeuGlnGlyCysSerThrSerAlaGlyArgPro 634
1547 ATYTTTNGCNGCNYTNCARCCNMNGCNGCNGCNGCNGCNGCNG 1596
635 AlaAspArgValGluArgValThrGlnAlaLeuArgSerAlaLeuAspSe 651
1597 CARGAMGNCNCARCARCTMNSMNGCNYTCARCCNCNTNGATYMS 1646
651 rCys.....ThrSerSerGluAlaProGly 660
1647 NTATYTCAYCCNCNCNGCNGCNGCNGCNGCNGCNG 1680

seq_name: gb_hlg:AC024612
seq_documentation_block: 166620 bp DNA linear HTG 06-MAY-2001
LOCUS AC024612
DEFINITION Homo sapiens chromosome 3 clone RP11-481H17 map 3, WORKING DRAFT
SEQUENCE, 22 unordered pieces.
ACCESSION AC024612
VERSION AC024612.2 GI:9929740
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 166620)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 3, clone RP11-481H17
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 166620)
AUTHORS Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bede, F.,
Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Deatrelano, K., Dewar, K., Diaz, J.S.,
Dodgson, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Glend, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, D.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Lacombe, K., Lamazares, R., Landers, T., Lehotzky, J.,
Levine, R., Lieu, C., Liu, G., Lock, A., MacDonald, P., Margulis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,
Meldrum, J., Menus, L., Minova, T., Miranda, C., Mlewa, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessier, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, D., Zimmer, A. and Zody, M.
DIRECT SUBMISSION
TITLE JOURNAL
COMMENT Submitted (01-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 26, 2000 this sequence version replaced gi:7132995.
All repeats were identified using RepeatMasker.

```


Smt. A.F.A. & Green, P. (1996:1997)
<http://ftp.genome.washington.edu/RW/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center Project name: L1857

Center Clone name: 481_H_17

Summary Statistics

Sequencing vector: M13; M7815; 95% of reads

Sequencing vector: Plasmid; n/a; 0.1% of reads

5.03173463177115Chemistry: Dye-terminator Big Dye; 100% of

reads

Assembly program: Phrap; version 0.960731

Consensus quality: 15636 bases at least Q40

Consensus quality: 161786 bases at least Q30

Consensus quality: 163606 bases at least Q20

Insert size: 156000; agarose-fp

Insert size: 164520; sum-of-contigs

Quality coverage: 4.9 in Q20 bases; agarose-fp

Quality co.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 22 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 5554: contig of 5554 bp in length
5555 5654: gap of 100 bp
5655 6844: contig of 1190 bp in length
6845 6944: gap of 100 bp
6945 8557: contig of 1613 bp in length
8558 8657: gap of 100 bp
8658 11247: contig of 2590 bp in length
11248 11347: gap of 100 bp
11348 13427: contig of 2080 bp in length
13428 13527: gap of 100 bp
13528 17090: contig of 3563 bp in length
17091 17190: gap of 100 bp
17191 20876: contig of 3686 bp in length
20877 20976: gap of 100 bp
20977 25280: contig of 4304 bp in length
25281 25380: gap of 100 bp
25381 29566: contig of 4186 bp in length
29567 29666: gap of 100 bp
29667 33425: contig of 3759 bp in length
33426 33525: gap of 100 bp
33526 37279: contig of 3754 bp in length
37280 37379: gap of 100 bp
37380 44779: contig of 7400 bp in length
44780 44879: gap of 100 bp
44880 52510: contig of 7631 bp in length
52511 52610: gap of 100 bp
52611 59797: contig of 7187 bp in length
59798 59897: gap of 100 bp
59898 66773: contig of 6876 bp in length
66774 66873: gap of 100 bp
66874 76191: contig of 9318 bp in length
76192 76291: gap of 100 bp
76292 100510: contig of 24219 bp in length
100511 100610: gap of 100 bp
100611 111192: contig of 10582 bp in length
111193 111292: gap of 100 bp
111293 124816: contig of 13524 bp in length
124817 124916: gap of 100 bp
124917 142057: contig of 17141 bp in length
142058 142157: gap of 100 bp
142158 157943: contig of 15786 bp in length
157944 158043: gap of 100 bp

```

FEATURES * 158044 166620: contig of 8577 bp in length.
 source location/Qualifiers

1. 166620
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /map="3"
 /clone="RP11-481H17"
 /clone_1bp="RP11-11 Human Male BAC"
 1. 5554
 /note="assembly-fragment"
 clone_end:SP6
 vector_side:left
 5655 .6844
 /note="assembly-fragment"
 6945 .8557
 /note="assembly-fragment"
 8658 .11247
 /note="assembly-fragment"
 11348 .13427
 /note="assembly-fragment"
 13528 .17090
 /note="assembly-fragment"
 17191 .20876
 /note="assembly-fragment"
 20977 .25280
 /note="assembly-fragment"
 25381 .29566
 /note="assembly-fragment"
 29667 .33425
 /note="assembly-fragment"
 33526 .37279
 /note="assembly-fragment"
 37380 .44779
 /note="assembly-fragment"
 44880 .52510
 /note="assembly-fragment"
 52611 .59797
 /note="assembly-fragment"
 59898 .66773
 /note="assembly-fragment"
 66874 .76191
 /note="assembly-fragment"
 76292 .100510
 /note="assembly-fragment"
 100611 .111192
 /note="assembly-fragment"
 111293 .124816
 /note="assembly-fragment"
 124917 .142057
 /note="assembly-fragment"
 142158 .157943
 /note="assembly-fragment"
 158044 .166620
 /note="assembly-fragment"
 clone_end:477
 vector_side:right

BASE COUNT 42182 a 40159 c 39756 g 42488 t 2102 others

alignment_scores:
 Quality: 692.50 Length: 470
 Ratio: 2.663 Gaps: 14
 Percent Similarity: 55.319 Percent Identity: 39.574

alignment block:
 US-09-899-471-2 x AC024612

Align seg 1/1 to: AC024612 from: 1 to: 166620

298 ValThrLeuGlyTrpGlnAlaProAspGlnSerProGlyGlnProLeuVa 314
 ||| ::|||::: |||||:::|||||

430 AlAcysPrometasrlystytIleHisArgAlgtTrrvalLeuValTrpLe 446

99570	CCAGACATTCACAAAGCCCTGGCCCTGTGTGGCT	99536
446	uAla	ysLeuLeuLeuAlaAlaAlaLeuPhePheLeuLeuLeuLysL	463
99535	GGCGTGGCTACTCTTTGGCGCGTGGCGTTTCCCTCATCGCTCTTCACAAA	99486	
463	ysAspArgArgLys.....	467	
99485	AGCATACACCGAAAGGTAGAGCGCTTCCGGCTCCCATTCCTCGTGGGGA	99436	
467	467	
99435	GGACGAGACTGCTGTGGAGTTCTCCGAGGAAAGCGGCGCGAGCTC	99386	
467	467	
99385	ACTGGCTGCTCCGCCCTCTCCCTAACAGGGTGGCTAGGCTTTGAA	99336	
467	467	
99335	ACAGAGCGTCCGCTCGGGGGGTGAGTGGAGCAAGCGCTGGCGGAGGGC	99286	
467	467	
99285	CGCCCGGGGAGCCAGGCGTGTGCCAGTCACTCTTCTCCATCT	99236	
468	AlaAlaArgGlySerArgThrAlaLeuLeuLeuH	479
99235	GTTTCTCGGACGAGCGCGCGGAGGAGC...CGCGGGTGTGCTCTCT	99189	
479	IsSerAla	spLysAlaIValIaGlyTyrGluArgLeuValGlyAlaLeuAlaSer	495
99188	ACTCAGCGCATCATCTGGGTTTCGAGCGCTGTGGGCGCTCGGCGTTC	99139	
496	AlaLeuSerGlnMetProLeuArgValAlaValaAspLeuTTPSerArgAr	512	
99138	GCCCTGTGCACACTCCGCTGCGCGTGGCGCTGAGACTGTGGAGCGTGC	99089	
512	GluLeuSerAlaHisGlyAlaAlaLeuAlaTyrPheHisHisGlnArgArgA	529	
99088	TCAACTGACGCGCCAGGGGCCCGTGGCTTGGTTACGCGGACGCGCGCC	99039	
529	ArgLeuLeuGlnGlyIValValaIleLeuLeuPheSerProAlaAla	545	
99038	AGACCTCGAGGAGGCGCGGTGGTGTCTTCTCTCCGCGTGC	98989	
546	ValAlaGlnCysGlnGlnTyrLeuGlnLeuGlnThrValGluPro....	560	
98988	GTGGGCTGTGCAGCGAGTGGCTACAGGATGGGGTCCGGGCCGGGGGC	98939	
561GlyProHisAspAlaLeuAlaAlaTyrLeuSerCysValLeuPro	576	
98938	GCAGGCGCCGACGACGCGCTTCCGCGCTCGCTACGCTGCGTGGTGGCCG	98889	
576	spPheLeuGlnGlyArgAlaThrGlnArgTyrValGlyValGlyTyrPheAsp	592	
98888	ACTTCTTCAGGCGCGGGCGCGCCGACAGTACGTGGGGCGCGTTCGAC	98839	
593	GlyLeuLeuHisProAspSerValProSerProPheArgValAlaProLe	609	
98838	AGCGTGTCCACCCGAGCGCGTACCCGCCCTTTTCCGACCGTCCGCT	98789	
609	uPheSerLeuProThrGlnLeuProAlaPheLeuAspAlaLeuGlnGly	626	
98788	CTTCACTGCTCCCTCCCACTGCCAGACTTCTCTGGGGCCCTCGACAGCAGC	98739	
626	IlyCysSerThrSerAlaGlyArgProAlaAspArgValGlyArgValThr	642	
98738	CTCGGCGCCGCGTTCGGGGGCTCCCAAGAGAGCGGAGCAAGTGTCC	98689	
643	GlnAlaLeuArgSerAlaLeuAspSerCys.....ThrSerSerI	657	
	

Sun Sep 29 09:32:43 2002

us-09-899-471-2.rge

Page 25

98688 CGGGCCCTTCAGCCAGCCCTGGATAGTACTTCCATCCCCGGGGGACTCC 98639

657 uAlaProGly 660

|||||||

98638 CGGGCCGGGA 98629


```

1766 ATGCTTAGGCCCTGCACCCCTAGCCTCCAGCTGGAGTACACTTGGGGCTAAG 1815
      ::::: ||| ||| |||
635 AlAAsPArgValGluArgValThrGlnAla.....LeuArgse 647
      ::::: ||| ||| |||
1816 CGGTGCTTGAAGAAACCGCTGGAGCAAGTGCACCTGCTGGAACTTGAAGC 1865
      ::::: ||| ||| |||
647 rAlAlaLeuAspSerCysThrSerSerGluAlaPro 659
      ||| ||| ||| |||
1866 TGGCAAGAGATGACTACCAAGGCGTCAACCATATGCCCC 1902
      ||| ||| ||| |||
Seq_name: /cgn2_6/ptodate/2/ina/5B_COMB.seq:US-08-620-654A-1

```

```

seq_name: /cgn2_6/pltdata7/ina/5b_COMB.seq:us-08-620
seq documentation block:
  Sequence 1, Application US/08620694A
  Patent No. 5869286
  GENERAL INFORMATION:
    APPLICANT: Yao, Zhengbin
    APPLICANT: Spriggs, Melanie
    APPLICANT: Fasllov, William
    TITLE OF INVENTION: No. 5869286el Receptor That I
    NUMBER OF SEQUENCES: 10
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Immune Corporation
      STREET: 51 University Street
      CITY: Seattle
      STATE: WA
      COUNTRY: USA
      ZIP: 98101
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: Apple Power Macintosh
      OPERATING SYSTEM: Apple Operating System 7.5.5
      SOFTWARE: Microsoft Word for Apple, Version 6.0
      CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/620,694A
        FILING DATE: 21 MARCH 1996
        CLASSIFICATION: 435
      PRIOR APPLICATION DATA:
        APPLICATION NUMBER: USSN 08/538,765
        FILING DATE: 7 AUGUST 1995
        CLASSIFICATION: 435
      PRIOR APPLICATION DATA:
        APPLICATION NUMBER: USSN 08/410,535
        FILING DATE: 23 MARCH 1995
        CLASSIFICATION: 435
      ATTORNEY/AGENT INFORMATION:
        NAME: Perkins, Patricia Anne
        REGISTRATION NUMBER: 34,695
        REFERENCE/DOCKET NUMBER: 2617-B
        TELECOMMUNICATION INFORMATION:
          TELEPHONE: (206)587-0430
          TELEFAX: (206)
      INFORMATION FOR SEQ ID NO: 1:
        SEQUENCE CHARACTERISTICS:
          LENGTH: 3288 base pairs
          TYPE: nucleic acid
          STRANDEDNESS: single
          TOPOLOGY: linear
        MOLECULE TYPE: cDNA to mRNA
        HYPOTHETICAL: NO
        ANTI-SENSE: NO
        ORIGINAL SOURCE:
          ORGANISM: Mouse
          STRAIN: HVS13 receptor
          FEATURE:
            NAME/KEY: CDS
            LOCATION: 121..2715
            US-08-620-694A-1
  Alignment_scores:
    Quality: 174..50
    Length: 582

```


Ratio: 0.687 Gaps: 28
Percent Similarity: 43.643 Percent Identity: 20.619

Alignment block:

US-09-899-471-2 x US-08-620-694A-1 ..

Align seg 1/1 to: US-08-620-694A-1 from: 1 to: 3288

```

224 LysAsnLeuThr..... GlyProGlnAsnIleThrLeuAsn..... 235
|||||
316 AAAAACCTGACCCCGCTTCCCAAAAAACATCTATATCAATCTTACTGT 365
236 .....HisThrAspLeuValProCysLeuCysIleGlnValT 248
|||
366 TTCCTCTACCCAGACGAGGAAATTAATCCCTGCTGTGCAATGTTGAG...T 412
248 TPSeuLeuGluProAsp..... 253
|||||
413 GGACCCCTGCAGACAGATGCGACATCTGTACCTCGAGGGTCAGAGCTG 462
254 .....SerGluArg..... ValGluPheCysPr 261
463 TCCTGCTCGACGTCGACACACATGACGGCTGTGTGTCAAGTTC...CA 509
261 oPheArgGluAspProGlyAlaHisArgAsnLeuThrPheHisIleAlaArgL 278
|||
510 GTTCTGTCCATGCTGCAGCATCACCGTAGAGCGGTGCTTCTCTCA 559
278 euArg... ValLeuSerProGlyValTPePLeuAspAlaProCysCys 293
|||
560 GCCACTTGTGTGTAGATCTCTGC..... 582
294 LeuProGlyLysValThrLeuGlyStrpGlnAlaProAspGlnSerProCys 310
583 CAGAGAGTGAAGTACTGTT..... 603
310 sGlnProLeuValProProValProGlnLysAsn..AlaThrValAsnGlu 326
|||
604 .CACCACTCTGCCGAGCCCATCTGTGATGGGAGCCCAACCAACAAATCCA 652
327 ProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuGlyValGluVal 343
653 .....AGATCATCTTGTGCTGCTGACTGT 675
343 LserThr..... 345
676 GAGACAGACAGATGAATGACTACCTCATCGTGAGCTCAGGACGACT 725
346 .....TrpGluLysValGlnLeuGlnAlaCysSer 355
726 TTGGGATCCCAACATCATCTGTGAGACCTTGACACACACAGCATCTGCGAG 775
356 TrpAlaAspSerLeuGlyProPheLysAspAspMetLeuValGluMet 372
776 TGG..... 778
372 tLysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluPro..... 386
779 .....ACTTCACCCCTGTGTGAATGAATCCACCCCTGA 809
387 .SerGlyCysThrProLeu...ProSerMetalSerThrArgAlaAla 401
810 CCAGGTCTCTGCGAAAGTTTCTCCGACTCAGACAGAACACAGCTGCTTGG 859
402 ArgLeuGlyGluGlnLeuLeuGlnAspPheArgSerHisGlnCysMetG 418
860 ATGTCGTAACCAATATTTGGCCAGCCAGAGAAATTCATCCATCAGCGA 909
418 nLeuTrpAsnAspAspAsnMetGlySerLeu..TrpAlaCys..... 431
910 GCTATATGTCACATTCATCTTAAGCAAGTTTCACTGTGCTGCCATCAGCA 959
431 ..... 431

```

```

960 CGTGCAGTTCACACCCCTTCTTCAAGACGCTGCCATAATGACTGTTTGAGAC 1009
431 ..... 431
1010 ACGCTGTGACTGTGCCCTGCCAGTAATCTCAATACACAGATCCCAAG 1059
432 PrometAspLysTrpIleHisArgArgTrpVal.....LeuValTrpLe 446
1060 CCAGTTGCAGACTACAT...CCCTGTGGGTGTATGGCCCTCATACACT 1106
446 uAlaCysLeuLeuAlaAlaLeuPhePhePheLeu..... 459
1107 CATGCGCATTTCTGTGTGGATCTGTCAATCGCTGATCATCTGTATGTA 1156
460 .....LeuLeuLysLysAspArgArgLysAlaAlaArgLysSer... 472
1157 CCTGAGAGCTTCTGTGGCGGCGATCAAGAAACATGATGATGACTCCAA 1206
472 ..... 472
1207 ATCAATGGCATCTTGCCCTGACAGACCTGATCCCCACCCCTGAGGCC 1256
473 .ArgThrAlaLeuLeuLeuHisSerAlaAspGlyAlaGlyTrpGluArgL 489
1257 CAGGAAGTCTGTGATGCTTACTGTGGCGGACACACCCCTTATGTGGAGG 1306
489 euValGlyAlaLeuAlaSerAlaLeu...SerGlnMetProLeuArgVal 504
1307 TGTCTCTAAGTTTCCGCCAGTTCCTGATCACTGCTGTCGACATGAGTA 1356
505 AlaValAspLeuTrpSerArgArgGluLeuSerAlaHisGlyAlaLeuAl 521
1357 GCCCTTGACCTCTCTGAGAGAGCAGGTATCTCTGAGGTGGGTGATGAC 1406
521 aTrpPheHisGlnArgArgArgLLeuGlnGlnGluValValVal 538
1407 CTGGGTGAGCCACACAGACAGAGATGTTGAGAGACACTCCAAATCA 1456
538 LeuLeuPheSer..... 542
1457 TCATCTCTGTGTCCGAGCAGCCACCAAGAGGAAAGTATCTTGCGT 1506
543 .....ProAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnLeuG 556
1507 TGGGCTGAGCGCTGTGTCCAGCTACGCTGAGCCACTGG..... 1545
556 nThrValGluProGlyProHisAspAlaLeuAlaIleTrpLeuSerCysV 573
1546 .....AGGCTGCT...GGGACCTTTTCACTGCAGCCATGAATGA 1585
573 aLeuProAspPheLeuGlnGlnArgAlaThrGlyArgTrpValGlyVal 589
1586 TCCGTGCAGACTTCAAGAGCCAGCCCTGTGGCAGCTGATGTTTGGC 1635
590 TyrPheAspGlyLeuLeuHisProAspSerValProSerProPheArgVa 606
1636 TACTTCAGTGCATCTGTAGTGAAGAGGATGTCGAGCTCTTCAACAT 1685
606 lAlaProLeuPheSerLeuPro..ThrGlnLeuProAlaPheLeuAspAla 622
1686 CACCTCCAGAGTACCACTCATGACAGATTGAGAGGTTTACTCTCGGA 1735
623 LeuGln.....GlyGlyCysSer.....Th 629
1736 TCCAGAGACTGAGATGTTGAACCGCGGAGTGCACCATGTCAGAGAG 1785
629 rSerAlaGlyArgProAlaAspArgValGluArgValThrGlnAlaLeuAl 646
1786 CTCACAGGGCAATTAATCTGCAGAGCCCTAGTGGCGGAGCTCAAGGA 1835
646 rGSerAlaLeuAspSerCysTrpSerSerSerGluAlaPro 659
|||

```


1836 GGCTGTGCTTAGGTTCCAGAGTGCAAAACCCAGTCCCG 1876
seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:us-09-022-255-1

seq_documentation_block:

Sequence 1, Application US/09022255
Patent No. 6072033

GENERAL INFORMATION:

APPLICANT: Yao, Zhenbin
APPLICANT: Spriggs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,255

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/620,694
FILING DATE: 21 MARCH 1996
APPLICATION NUMBER: USSN 08/538,765
FILING DATE: 7 AUGUST 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 3288 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse
STRAIN: HVS13 receptor
FEATURE:
NAME/KEY: CDS
LOCATION: 121..2715
US-09-022-255-1

alignment_scores:

Quality: 174.50 Length: 582
Ratio: 0.687 Gaps: 28
Percent Similarity: 43.643 Percent Identity: 20.619

alignment_block:

US-09-899-471-2 x US-09-022-255-1

Align seg 1/1 to: US-09-022-255-1 from: 1 to: 3288

224 LysAsnLeuThr.....GlyProGlnAsnIleThrLeuAsn..... 235

|||||.....
316 AAAAAGCTGACCCCGTCTTCCCAAAAAACATCATATCATCTAGT 365
236HisThrAspLeuValProCysIleGlnValT 248
366 TTCTCTTACCCAGACGAGAAATTAGCTGTTGCATGTTGAG...T 412
248 rPserLeuLuproAsp..... 253
413 GGACCTGCAGACAGATGCCAGCATCTGACTCGAGGTCGAGACCTG 462
254SerGluArg.....ValGluPheCysPr 261
463 TCCGCTCTCAGCTGACGACCAATGAGCGGCTGTGTCAAGTTC...CA 509
261 oPheArgIuAspProGlyAlaHisArgAsnLeuThrPheIleAlaArgL 278
510 GTTCTGTGCTGCTGCTGACGATCAGCGGTGCGGTTTCTCTCA 559
278 euArg...ValLeuSerProGlyValTTrpIleLeuAspAlaProCysCys 293
560 GCCACTTGTGTGATAGATCCTGCC..... 582
294 LeuProGlyLysValThrLeuCysTrpGlnAlaProAspIleSerProCys 310
583 CAGGATGATGAAGTACTGTT..... 603
310 sGlnProLeuValProProValProGlnLysAsnAlaThrValAsnGlu 326
604 CACCACTGCGGAGCCATCCCTGATGGGAGCCCAACCAACCAATCCA 652
327 ProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuCysValGlnVal 343
653AGATCATCTTGTGCTGCTGACTGT 675
343 lserThr..... 345
676 GAGGACAGCAATGATGATGACTACTCATGCGTGAGCTGAGCAACCT 725
346TrpGluLysValGlnLeuGlnAlaCysSer 355
726 TTGGATCCCAATCATCATGCTGAGAGACTTGGACACACATCTGGAG 775
356 TrpAlaAspSerLeuGlyProPheLysAspMetLeuLeuValGluMet 372
776 TGG..... 778
372 llyThrGlyLeuAsnAsnThrSerValCysAlaLeuGluPro..... 386
779ACTTCACCTGTGTGATGATGATCCACCCCTCA 809
387 ..SerGlyCysThrProLeu...ProSerMetAlaSerThrArgAlaAla 401
810 CCAGGTCTCTGTAAGATTTCTCCGACTCAGACAGACAGCTGTTTG 859
402 ArgLeuGlyGluGlnLeuLeuGlnAspPheArgSerHisGlnCysMetG 418
860 ATGTGCTTAACAAATATTGTCGCGCAGGCAAGAAATTCTCATACGCA 909
418 nLeuTrpAsnAspAspAsnMetGlySerLeu...TrpAlaCys..... 431
910 GCTATGTGCATTCATCTAAGCAAGTTTCACTGCTGCTGCATACCA 959
431 431
960 CGTGCAGCTCCAGCCCTTCTTCCAGAGCTGCTAATGACTGTTGAGAC 1009
431 431
1010 ACGCTGACTGTGCTGCCCGCAGTAATCTCAAAATACACAGTTCACAG 1059
432 ProMetAspLysTrpIleHisArgArgTrpVal.....LeuValTrpLe 446
|||||.....

413 GGACCTGCAGACAGATGCCAGCATCTGTACTGAGAGGTGCAGAGCTG 462
 254 SerGluArg..... ValGluPheCysPr 261
 463 TCCGCTCTGAGCTGACCAACCAAGCGCTGTGTGTCAGATTC...CA 509
 261 oPheArgGluAspProGlyAlaHisArgAsnLeuTrpHisIleAlaArgL 278
 510 GTTTCGTGCATCTGCAGCATCACCCTTAACGGGTGGCGTTTCTTCA 559
 278 euArg...ValLeuSerProGlyValTrpGluLeuAspAlaProCysCys 293
 560 GCCACTTGTGTAGATCCCTGGC..... 582
 294 LeuProGlyLysValThrLeuCysTrpGlnAlaProAspGlnSerProCys 310
 583 CAGAGTATGAGTACTGTT..... 603
 310 sGlnProLeuValProProValProGlnLysAsn..AlaThrValAsnGlu 326
 604 .CACCACCTGCCAGACCCATCCCTGTATGGGGACCCAAACCAATCCA 652
 327 ProGlnAspPheGlnLeuValAlaGlnHisProAsnLeuCysValGlnVa 343
 653 AGATCATCTTGTGCTGACTGT 675
 343 LserThr..... 345
 676 GAGACAGACAGATGAGATGACTACTGCTGAGCTGAGCTGAGCGAGCT 725
 346 TrpGluLysValGlnLeuGlnAlaCysSer 355
 726 TTGGATGCCACATCATCTGTGGAGCTTGAGACACACAGCATCTGGCAG 775
 356 TrpAlaAspSerLeuGlyProPheLysAspPheLeuValGluMe 372
 776 TGG..... 778
 372 LysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluPro..... 386
 779 ACTTCACCTGTGGAATGAAATCCACCCCTA 809
 387 ..SerGlyCysThrProLeu...ProSerMetAlaSerThrArgAlaAla 401
 810 CCAGGCTCGCTGGAAGATTCTCCGACTCAGAGAACACACAGCTGCTTGG 859
 402 ArgLeuGlyGluGluLeuLeuGlnAspPheArgSerHisGlnCysMetG1 418
 860 ATGTCCTTAAACAATATTGCGCCCGCAGCAAGAAATTCATCATCAGCGA 909
 418 nLeuTrpAsnAspAspAsnMetGlySerLeu..TrpAlaCys..... 431
 910 GCTAATGTCATTCATCTAAGCAAGTTTCAGTGTGCTCCATCAGCA 959
 431 431
 960 CGTGACAGTCCACCCCTTCTTACAGACACTGCTAATGACTGTTGAGAC 1009
 431 431
 1010 ACGCTGTGACTGTGCCCTGCCAGTAATCTCAATATACCACAGTCCCAAG 1059
 432 PrometAspLysTrpIleHisArgArgTrpVal.....LeuValTrpLe 446
 1060 CCAAGTTCAGACTACAT...CCCTGTGTGGTATGAGCCCTCATCAGACT 1106
 446 uAlaCysLeuLeuAlaAlaAlaLeuPhePheLeu..... 459
 1107 CATGCCATTTCTGTGTGGATCTGTCAATCGTCTGATCATCTGTATGA 1156
 460LeuLeuLysLysAspArgArgLysAlaAlaArgGlySer... 472
 1157 CCTGAGAGCTTCTGTGGCGCATCAAGAAACATGTGATGACTCCAAA 1206

472 472
 1207 ATCAATGGCATCTTCCCGTAGACAGCTGACTCCCAACCCCTGAGCC 1256
 473 .ArgThrAlaLeuLeuLeuHisSerAlaAspGlyAlaGlyTrpGluArgL 489
 1257 CAGGAAGTCTGATGCTGTACTGCGCCGACCAACCCCTTATGTGGAGG 1306
 489 euValGlyAlaLeuAlaSerAlaLeu...SerGlnMetProLeuArgVal 504
 1307 TGGTCTTAAAGTTGCCAGTCTGCTGATCATCTGCTGTGGACGTAGA 1356
 505 AlaValAspLeuTrpSerArgArgGluLeuSerAlaHisGlyAlaLeuAl 521
 1357 GCCCTTACCTCTCGAAGACAGTATCTGTAGGTGGGGGTATGANC 1406
 521 aTrpPheHisGlnGlnArgArgGlyLeuGlnGluGlyValValI 538
 1407 CTGGGTAGCCGACAGAAAGCAGATGTGTGAGAGCAATCCAAATCA 1456
 538 LeuLeuPheSer..... 542
 1457 TCATCTGTGTCCCGAGCGACCCAGCAAGTGAAGTATCTTGGGT 1506
 543ProAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnLeuG1 556
 1507 TGGGCTGAGCTGTCTGTCCAGCTTACGCTGAGTGCACCTGG..... 1545
 556 nThrValGluProGlyProHisAspAlaLeuAlaThrLeuSerCysV 573
 1546AAGCTGCT...GGGACCTTTTCAGTCGACCAACCAAGCA 1585
 573 aLeuProAspPheLeuGlnGlyArgAlaThrGlyArgTrpValGlyVal 589
 1586 TCTGGCAGACTTCAAGAGGCCAGCTGCTCGGACCTAGCTTGTTCG 1635
 590 TyrPheAspGlyLeuLeuHisProAspSerValProSerProPheArgVa 606
 1636 TACTGTGGCATCTGTAGTGAAGGATGTCCCGACCTTTCACAT 1685
 606 lAlaProLeuPheSerLeuPro..ThrGlnLeuProAlaPheLeuAspAla 622
 1686 CACCTCAGGATCCCACTCATGACAGATTGAGAGGTTTACTTCGGA 1735
 623 LeuGln.....GlyLysCysSer.....Th 629
 1736 TCCAGGACTGAGATGTTGAACCCGCGGAGATGCACCATGTGAGAG 1785
 629 rSerAlaGlyArgProAlaAspArgValGluArgValThrGlnAlaLeu 646
 1786 CTCACAGGGACAAATTACTCTGCAAGCCCTTAGTGCCGCGAGCTCA 1835
 646 rSerAlaLeuAspSerCysThrSerSerSerGlnAlaPro 659
 1836 GGCTGTCTTAGTTCCAGAGATGGCAACCAACCAAGTCCCG 1876
 seq.name: /cgn2L6/ptodata/2/lna/6A.COMB.seq: us-08-978-773-1
 seq.documentation block:
 : Sequence 1, Application us/08978773
 : Patent No. 6083906
 : GENERAL INFORMATION:
 : APPLICANT: Trout, Anthony
 : TITLE OF INVENTION: Method of Regulating Nitric Oxide Production
 : NUMBER OF SEQUENCES: 4
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Immunex Corporation
 : STREET: 51 University Street
 : CITY: Seattle
 : STATE: WA
 : COUNTRY: USA
 : ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: Apple PowerMacintosh
 OPERATING SYSTEM: Apple Operating System 7.5.5
 SOFTWARE: Microsoft Word for PowerMacintosh, Version 6.0.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/978,773

FILING DATE:

CLASSIFICATION: 530
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 60/052,525
 FILING DATE: 27 NOVEMBER 1996

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia Anne
 REGISTRATION NUMBER: 34,693
 REFERENCE/DOCKET NUMBER: 2623-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206)587-0430
 TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3288 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Mouse

IMMEDIATE SOURCE:

CLONE: IL-17 receptor

FEATURE:

NAME/KEY: CDS

LOCATION: 121..2712

US-08-978-773-1

Alignment scores:

Quality: 174.50 Length: 582

Ratio: 0.687 Gaps: 28

Percent Similarity: 43.643 Percent Identity: 20.619

alignment block:

US-09-899-471-2 x US-08-978-773-1

Align seg 1/1 to: US-08-978-773-1 from: 1 to: 3288

224 LysAsnLeuThr.....GLYProGlnAsnIlePheLeuAsn..... 235
 |||||
 316 AAAAAGCTGACCCGCTTCCCAAAAACATCTATATCAATCTTAGTGT 365
 236HisThrAspLeuValProCysLeuCysIleGlnValT 248
 |||||
 366 TTCCTTACCCACGACGAGAAATTAATGCTGCTGATCTTGAAG..T 412
 248 rPserLeuGluProAsp..... 253
 |||||
 413 GGACCTCGAGACAGATGCCAGCATCCTGTAACGAGAGGTGACAGCTG 462
 254SerGluArg.....ValGluPheCysPr 261
 |||||
 463 TCCGCTCTGACAGCTGAACACCAATGAGCGCTGTCTCAAGTTC...CA 509
 261 oPheArgGluAspProGlyAlaHisArgAsnLeuTrpHisIleAlaArgL 278
 |||||
 510 GTTTCGTGCTCAGCTGACAGATCAACGCTAAGCGGTTCCTTCCCTTCA 559
 278 euArg...ValLeuSerProGlyValTrpGlnLeuAspAlaProCysCys 293
 |||||
 560 GCCACTTGTGTGATGATCCTGGC..... 582

294 LeuProGlyLysValThrLeuCysTrpGlnAlaProAspGlnSerProCys 310
 ::::|
 583 CAGAGATGATGAAGTACTGTT..... 603
 310 sGlnProLeuValProProValProGlnLysAsn..AlaThrValAsnGlu 326
 ::::|
 604 .CACCACCTGCCAAGCCCATCCCTGATGGAGACCAACCAACCAATTCGA 652
 327 ProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuCysValGlnVa 343
 ::::|
 653AGATCATCTTGTGCTGACTGT 675
 343 lSerThr..... 345
 |||||
 676 GAGACACAGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 725
 346TrpGluLysValGlnLeuGlnAlaCysSer 355
 |||||
 726 TTGGATGCCAATCATCATCTGAGACACCTTGACACACACATCTGCGAG 775
 356 TrpAlaAspSerLeuGlyProPheLysAspAspMetLeuValGluIle 372
 |||||
 776 TGG..... 778
 372 tLysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluPro..... 386
 |||||
 779ACTTCACCTGCTGATGATGATGATGATGATGATGATGATGATG 809
 387 .SerGlyCysThrProLeu...ProSerMetAlaSerThrAlaAla 401
 ::::|
 810 CCAGTCTCTGCTGGAAGTTTCTCGACTGACAGACACACAGCTGTTTG 859
 402 ArgLeuGlyGluLeuLeuGlnAspPheArgSerHisGlnCysMetG1 418
 ::::|
 860 ATGTCGTTAAACAAATATTTGCGCCAGCAGCAAGAAATTCATCGACGA 909
 418 nLeuTrpAsnAspAspAsnMetGlySerLeu..TrpAlaCys..... 431
 ::::|
 910 GCTAATGTACATTCATCTAAGCAAGTTTCACTGCTGCTGCCATCAGCA 959
 431 431
 960 CGTGCAGCTCCAGCCCTTCTTACAGACGCTGCCAATATGACTGTTGACAC 1009
 431 431
 1010 ACGCTGACTGTGCCCTGCCAGATATCTCAAAATACACAGACTTCCGANG 1059
 432 ProMetAspLysTrpIleHisArgArgTrpVal.....LeuValTrpLe 446
 |||||
 1060 CCACTTCCAGACTACATT...CCCTGTGGGTGATGAGCTTATCATCACT 1106
 446 uAlaCysLeuLeuLeuAlaAlaLeuPhePheLeu..... 459
 |||||
 1107 CATGCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGA 1156
 460LeuLeuLysLysAspArgLysAlaAlaArgLys... 472
 |||||
 1157 CTTGAGAGCTTTTGGGGCCGATCAAGAGAAACATGGATGATCTCCAAA 1206
 472 472
 1207 ATCAATGCACTCTTCCCTGACAGACCTGACTCCCAACCCCTGAGGCC 1256
 473 .ArgThrAlaLeuLeuLeuHisSerAlaAspGlyValGlyTrpGluArgL 489
 |||||
 1257 CAGGAGCTGTGATGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGAG 1306
 489 euValGlyAlaLeuAlaSerAlaLeu...SerGlnMetProLeuArgVal 504
 ::::|
 1307 TGGCTCTAAAGTTGCGCCAGTCTGATGATGATGATGATGATGATGATGAT 1356
 505 AlavalAspLeuTrpSerArgArgGluLeuSerAlaHisGlyAlaLeuAl 521


```

1357 GCGCTTGACCTCTCGAGAGACAGCTTATCTGTAGAGTGGGGTCATGAC 1406
521 atpPhenHisGlnArgArgArgIleLeuGlnGlnGlyValValI 538
1407 CTGGGTGAGCCGACAGAGAGAGATGGTGGAGCACTCCAAATCA 1456
538 leuLeuPheSer..... 542
1457 TCATCCGTGTCTCCGAGGACCCAGCAAAAGTGGAAAGCTATCTGGGT 1506
543 .....ProAlaValAlaGlnCysGlnGlnTrpLeuGlnLeuG 556
1507 TGGGCTGAGCTCTGCTGACCTAGCTGCTGACCACTG..... 1545
556 ntValGlnGlnProGlnProHisAspAlaLeuAlaAlaTrpLeuSerCysV 573
1546 .....AAGCTGTCT...GGGGACCTTTTCTACCTGACCCAGCAATGACATGA 1585
573 alleuProAspPheLeuGlnGlnArgAlaThrGlyArgTrpValGlyVal 589
1586 TCCGTCGACAGCTTCAAGAGGCCAGCTCTTGGCACTTACCTGTGTTGC 1635
590 TyrPheAspGlyLeuLeuHisProAspSerValProSerProPheArgVa 606
1636 TACTTCAGTGCATCTGTATGTAGAGAGGATGTCCCGACCTTTCACAT 1685
606 lAlaProLeuPheSerLeuPro.ThrGlnLeuProAlaPheLeuAspAla 622
1686 CACCTCCAGGATCCCACTCATGACAGATTGTGAGAGAGTTTACTTCCGGA 1735
623 leuGln.....GlyGlyCysSer.....Th 629
1736 TCCAGAGACCTGGAGATGTTTGAACCCGCGGATGACACCATCTCAGAGAG 1785
629 rSerAlaGlyArgProAlaAspArgValGlnArgValThrGlnAlaLeuA 646
1786 CTCACAGGGGACATTTACTCAGAGCCCTAGTGGCCGCGACCTCAAGGA 1835
646 rSerAlaLeuAspSerCysThrSerSerGluAlaPro 659
1836 GGCCTGCTTGTAGTTCCAGAGAGTGGCAAAACCAAGTCCCGC 1876
seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:us-09-022-253-1
seq_documentation_block:
Sequence 1, Application US/09022253
Patent No. 6096305
GENERAL INFORMATION:
APPLICANT: Yao, Zhengbin
APPLICANT: Fanliow, William
TITLE OF INVENTION: No. 6096305el Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: Apple Power Macintosh
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,253
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/620,694
FILING DATE: 21-MARCH-1996

```

```

APPLICATION NUMBER: USSN 08/538,765
FILING DATE: 7 AUGUST 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3288 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse
STRAIN: HVS13 receptor
FEATURE:
NAME/KEY: CDS
LOCATION: 121..2715
US-09-022-253-1

alignment_scores:
Quality: 174.50 Length: 582
Ratio: 0.687 Gaps: 28
Percent Similarity: 43.643 Percent Identity: 20.619

alignment_block:
US-09-899-471-2 x US-09-022-253-1 ..
Align seg 1/1 to: US-09-022-253-1 from: 1. to: 3288

224 LysAsnLeuThr.....GlyProGlnAsnIleThrLeuAsn..... 235
316 AAAAACTGACCCCGCTCTTCCCAAAAAATCATATCAATCTTAGTGT 365
236 .....HistHAspLeuValProCysLeuCysIleGlnValT 248
366 TTCCTTACCCAGCAGCAGAGATTAAGTCCCTGTGTGCAATGTGAG...T 412
248 rPserLeuGluProAsp..... 253
413 GGACCTGTCAGACAGATGCCAGCATCTGTACTCGAGGCTCAGAGCTG 462
254 .....SerGluArg.....ValGluPheCysPr 261
463 TCCGTCCTGACAGCTGACACCAATAGCGGCTGTGTGCAAGTTC...CA 509
261 oPheArgGlnAspProGlyAlaHisArgAsnLeuThrHisIleAlaArg 278
510 GTTCTGTGCATGCTGCGACATCACCCTGTAAGCGGTGGCGGTTTCCTTCA 559
278 euArg...ValLeuSerProGlyValTrpGlnLeuAspAlaProCysCys 293
560 GCCACTTGTGTGATGATCTGGC..... 582
294 leuProGlyLysValThrLeuCysTrpGlnAlaProAspGlnSerProCys 310
583 CAGAGATATGAAAGTACTGTT..... 603
310 sGlnProLeuValProProValProGlnLysAsn..AlaThrValAsnGlu 326
604 .CACCACTGCGCAAGCAAGCCATCCGTGATGGAGGCCAAACCAACAATCA 652

```



```

327 ProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuCysValGlnVal 343
653 .....ACATCATCTTTGTGCTGACTGT 675
343 LserThr..... 345
676 GAGGACGACCAAGATGAGATGACTACCTGCGTGAGCTCAGGACGCT 725
346 .....TTPGluLysValGlnLeuGlnAlaCysSer 355
726 TTGGATCCCAACATCATCTGCTGAGACCTTGACACACACATCTTGAG 775
356 TrpAlaAspSerLeuGlyProPheLysAspAspMetLeuValGlnVal 372
776 TGG..... 778
372 TlyThrGlyLeuAsnAsnThrSerValCysAlaLeuGluPro..... 386
779 .....ACTTCACTGCTGAGATGATGATCCACCCCTTA 809
387 ..SerGlyCysThrProLeu...ProSerMetAlaSerThrArgAlaAla 401
810 CCAGGCTCTGCGAAGATTCTCCGACCTCAGAGAACACACAGCTGCTTGG 859
402 ArgLeuGlyGluGlnLeuLeuGlnAspPheArgSerHisGlnCysMetG 418
860 ATGCGCTTAACAATATTGCGCCCGACGACCAAGAACATTCATCAGCGA 909
418 nLeuTrpAsnAspAspAspMetGlySerLeu...TrpAlaCys..... 431
910 GCTAATGTCACATTCATCTAGCAAGTTTCACTGCTGCGCATCAGCA 959
431 ..... 431
960 CGTCAGGCTCCAGCCCTTCTTCAGAGCTGCTTAATGACTGTTTGAGAC 1009
431 ..... 431
1010 ACGGTGACTGTGCTCCCTGCCAGATATCTCAATATCCACAGTCCCAAG 1059
432 ProMetCysLysTrpIleHisArgArgTrpVal.....LeuValTrpLe 446
1060 CCAGTGTGAGACTACAT...CCCTGTGGGTATGAGCTTCATCAGCT 1106
446 AlaCysLeuLeuLeuAlaAlaLeuPhePheLeu..... 459
1107 CATGCGCATTTGCTGTGGGATGCTGATCGTGCATCATCTGTATGA 1156
460 .....LeuLeuLysLysAspArgLysAlaAlaArgLysSer... 472
1157 CCGGAGGCTTTTGGCGCCCATCAGAGAACATGATGACTCCAA 1206
472 ..... 472
1207 ATCAATGCGATCTTGGCCGTAGCAGACTGACTCCCGCCAGCTGAGGCG 1256
473 ..ArgThrAlaLeuLeuHisSerAlaAspLysAlaGlyTrpGluArg 489
1257 CAGGAGGCTGATGATCTGCTGCGCGACACCCCTCTATGTGAGG 1306
489 euValGlyAlaLeuAlaSerAlaLeu...SerGlnMetProLeuArgVal 504
1307 TGTGCTTAAGTTTCCCGAGTTCCGTGATCAGCTGCTGCGACTGAGTA 1356
505 AlaValAspLeuTrpSerArgArgLysLeuSerAlaHisGlyAlaLeuAl 521
1357 GCGCTTGACCTCTGAGAGAGAGAGAGGTATCTCTGAGGTGGGGTATGAC 1406
521 atTrpHisHisGlnArgArgArgLysLeuGlnGlnGlyLysValVal 538
1407 CTGGGTGAGCCGACAGAGAGAGATGATGATGAGCAACTCCAAATCA 1456

```

```

538 LeuLeuPheSer..... 542
1457 TCATCTGTGTTCGAGGACGACCAAGAACATGAAAGTACTTGGGT 1506
543 .....ProAlaAlaValAlaGlnCysGlnGlnTrpLeuGln 1545
1507 TGGGCTGAGCCCTGCTCCAGCTACGCTGTGACCACTGG..... 1545
556 nThrValGluProGlyProHisAspAlaLeuAlaAlaTrpLeuSerCys 573
1546 .....AAGCTGCT...GGGACCTTTTCACTGACGACCATGAACATGA 1585
573 alleuProAspPheLeuGlnGlyArgAlaThrGlyArgValGlyVal 589
1586 TCCTGCGAGACTTCAGAGGCGACGCTGCTGCGACCTAGCTTGTTC 1635
590 TyrPheAspGlyLeuLeuHisProAspSerValProSerProPheArgVa 606
1636 TACTTCAGTGGATCTGTAGTGAAGGATGTCGCCACCTTTCACAT 1685
606 AlaProLeuPheSerLeuPro...ThrGlnLeuProAlaPheLeuAspAla 622
1686 CACCTCCAGTACCCACTCATGAGATTTGAGAGAGTTACTTCCGGA 1735
623 LeuGln.....GlyGlyCysSer.....Th 629
1736 TCCAGACCTGAGATGTTGAACCCGCGGATGACCATGTGACAGAG 1785
629 rSerAlaGlyArgProAlaAspArgValGluArgValThrGlnAlaLeu 646
1786 CTCACAGGAGACATTAACCTGCGAGAGCCCTAGTGGCGGACGCTCAAGA 1835
646 rGSerAlaLeuAspSerCysThrSerSerSerGluAlaPro 659
1836 GCGTCTGCTTACGTTCCAGAGAGTGCAACCCAGTGGCCCG 1876

```

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-09-022-260-1

```

seq_documentation_block:
; Sequence 1, Application US/09022260
; Patent No. 6100235
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; TITLE OF INVENTION: No. 6100235el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,260
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/620,694
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/410,535
; FILING DATE: 23 MARCH 1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne

```


[illegible]


```

1546 .....:AGGCTGCT...GGGACCTTTTCACCTGACCATGAAACATGA 1585
573 alleuprpsphleuInglYarglathrglyAargTyValglYval 589
1586 TCTGTCCACACTTCAAGAGCCAGCCGCTTCGCGACCTGACCTGCTTTCG 1635
590 TyRphespGlYleuInHISProAspSerValProSerProPheargVa 606
1636 TACTTCAGTGGCATCTGTACTGACAGAGGATGTCCCGACCTTCAACAT 1685
606 lAlaProlleuSerleuPro-ThrgInleuProAlaPheleuAspAla 622
1686 CACCTCCAGTACCCACTCATGACATGATTGACAGAGGTTTACTCCGGA 1735
623 LeuGln.....GlyGlycysSer.....Th 629
1736 TCCAGGACCTGGAGATGTTGAACCCGCGGATGCACCATGTGACAGAG 1785
629 rSerAlaGlyArgProAlaAspArgValGluArgValThrgInAlaLeuA 646
1786 CTCACAGGGGACAAATTACCTGCAGAGCCCTAGTGCCGCGACGTCACAGA 1835
646 rGSerAlaLeuAspSerCysThrSerSerSerGluAlaPro 659
1836 GCGTGTGCTTAGGCTTCAGAGGTGCAGAAACCCAGTCCCGC 1876

Seq_name: /cgr2_6/ptodata/2/lna/6B_COMB.seq:US-09-022-259-1
Seq documentation_block:
Sequence 1, Application US/09022259
Patent No. 6191104
GENERAL INFORMATION:
APPLICANT: Yao, Zhenpin
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6191104el Receptor That Blinds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,259
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perlus, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3288 base pairs
TYPE: nucleic acid

```

```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse
STRAIN: HVS13 receptor
FEATURE:
NAME/KEY: CDS
LOCATION: 121..2715
US-09-022-259-1

alignment_scores:
Quality: 174.50 Length: 582
Ratio: 0.687 Gaps: 28
Percent Similarity: 43.643 Percent Identity: 20.619

alignment_block:
US-09-899-471-2 x US-09-022-259-1
Align seg 1/1 to: US-09-022-259-1 from: 1 to: 3288

224 lysasnleuThr.....GlyProGlnasnleuThrleuAsn..... 235
316 AANAACCTGACCCCGCTCTCCCAAAAACATATATCAATCTTAGTGT 365
236 .....HsthrAspLeuValProCysleuCysIleGlnValT 248
366 TTTCCTACCCAGCAGGAGAAATTAGTCCCTGTGTTCATGTGAG...T 412
248 rPSerleuGlnProAsp..... 253
413 GCACCTTCACAGACAGATGCCAGATCTGTACTGACAGGTGCAGAGCTG 462
254 .....SerGluArg.....ValGluPheCysTr 261
463 TCCGCTCTGCAGCTGAAACACCAATGAGCGCTGTGTCAAGTTC...CA 509
261 opheArgGlnAspProGlyAlaAlaHisArgAsnleuThrPheIleAlaArgL 278
510 GTTCTGTCCATCTGTCAGCATCAACCGTAAGCGGTGGCGGTTTCTTCA 559
278 euArg.....ValleuSerProGlyValTTrpGlnleuAspAlaProCysCys 293
560 GCCACTTGTGTGATGATCTGCG..... 582
294 leuProGlyLysValThrleuCysTrpGlnAlaProAspGlnSerProCys 310
583 CAGCAGTATCAAGTACTGT..... 603
310 sGlnProleuValProProValProGlnLysAsnAlaThrValAsnGln 326
604 .CACCACTGGCGAAGCCATCCTGATGGAGGCCCAACCAATATCA 652
327 ProGlnAspPheGlnleuValAlaGlyHisProAsnleuGlyValGlnVal 343
653 .....AGATCAATCTTGTGCTGACTGT 675
343 lSerThr..... 345
676 GAGACACAGACAGATGAAGACTACTCATGCGTGAAGCTCAGCAGCGCT 725
346 .....TrpGlnLysValGlnleuGlnAlaCysSer 355
726 TTGGATCCCAACATCACTGTGGAGACCTTGACACACAGCATTCGCGAG 775
356 TrpAlaAspSerleuGlyProPheLysAspAspMetleuLeuValGluMe 372
776 TGG..... 778
372 tlyThrGlyLeuAsnAsnThrSerValCysAlaLeuGlnPro..... 386

```



```

779 .....ACTTCACCCCTGTGGAGATGATCCACCCCTTA 809
387 ..SerGlyCysThrProLeu...ProSerMetAlaSerThrArgAlaAla 401
810 CCAGGTCCTGCTGGAAAGTTTCTCCGACTCGAGAACCCACAGCTCTTGG 859
402 ArgLeuGlyGluGluLeuGlnAspPheArgSerHisGlnCysMetGlu 418
860 ARGTCCTTAAACAATATTGCGCCAGGACGAGAAAGAAATTCATCAGCGA 909
418 nLeuTrpAspAspAspAsnMetGlySerLeu...TrpAlaCys..... 431
910 GCTAAGTGTACATTCACCTCTAAGCAACTTTCACGTGCTCTCCATCACA 959
431 ..... 431
960 CGTGACAGTCCACCCCTTCTTTCAGACAGCTGCTTAATGATGTTTGAGAC 1009
431 ..... 431
1010 AGCGTGTGACTGTGCCCCGTGACGTAATCTCAATATCCACAGTTCCCAAG 1059
432 PrometAspLysPylIleHisArgArgTrpVal.....LeuValTrpLe 446
1060 CCAGTGTGACAGACTACATTT...CCCCGTGGGTGTATGGCTCATCAGACT 1106
446 uAlaCysLeuLeuLeuAlaAlaAlaLeuPhePheLeu..... 459
1107 CATGCCATTCCTGCTGGTGGATCTGTCATCGTGTGATCATCTGTATGA 1156
460 .....LeuLeuLysLysAspArgArgLysAlaAlaArgGlySer... 472
1157 CCTGGAGGCTTCTGTGGCGCCGATCAAGAAAGAAATGATGACTCCAAA 1206
472 ..... 472
1207 ATCAATGCAATCTTGCCCGTAGCAGACTGACTCCACCCCTGAGAGCC 1256
473 ..ArgThrAlaLeuLeuLeuHisSerAlaAspGlyAlaGlyTrpGluArgL 489
1257 CAGGAAGTCTGTGATCTGATCTGATCGGCCGACACCCCTCATGTGAGG 1306
489 euValGlyAlaLeuAlaSerAlaLeu...SerGlnMetProLeuArgVal 504
1307 TGGTCTTAAGTTCGCCAGCTGCTGATCTGATCTGCTGTGGCACTGAAGTA 1356
505 AlaValAspLeuTrpSerArgArgGluLeuSerAlaHisGlyAlaVal 521
1357 GCCCTTGACCTCTCTGAGAGAGCAGGTTATCTGAGGTGGGGGTGATGAC 1406
521 aTrpPheHisHisGlnArgArgArgIleLeuGlnGluGlyValVal 538
1407 CTGGGTGAGCCCGCAGAGAGAGATGTGTGAGAGCACTCCAAATCA 1456
538 LeuLeuPheSer..... 542
1457 TCAATCTGTGTTCCCGAGGACCCAAAGCAAGTGAAGCTATCTTGGGT 1506
543 .....ProAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnLeuGlu 556
1507 TGGGCTGAGCCCTGCTGTCAGTACGTTGTGACACATG..... 1545
556 nThrValGluTrpGlyProHisAspAlaLeuAlaAlaTrpLeuSerCysv 573
1546 .....AGCCCTGCT...GGGACCTTTCCTGCTGACGCACTCAACATGA 1585
573 aLeuProAspPheLeuGlnGlyArgAlaThrGlyArgTrpValGlyVal 589
1586 TCTCTGCAACTTCAAGAGCCGCTGCTTGGCAGCTTACGTTGTTGC 1635
590 TyrPheAspGlyLeuLeuHisProAspSerValProSerProPheArgVa 606

```

```

1636 TACTTCAGTGGCATCTGTAGTGAGAGGATGTCCCGACCTCTTCAACT 1685
606 lAlaProLeuPheSerLeuPro...ThrGlnLeuProAlaPheLeuAspAla 622
1686 CACTTCAGGATACCACACTCATGACAGATTTGAGAGGTTTACTTCCGGA 1735
623 LeuGln.....GlyGlyCysSer.....Th 629
1736 TCCAGGACCTTGAGATGTTTGAACCCGCGCGATGCACCATGTGACAGAG 1785
629 rSerAlaGlyArgProAlaAspArgValGluArgValThrGlnAlaLeu 646
1786 CTCACAGGGGCAATTTACTTCAGAGCCCTTACTGTGCGCCGACCTCAAGA 1835
646 rGSerAlaLeuAspSerCysThrSerSerSerGluAlaPro 659
1836 GCGTGTGCTTGGTTCCAGAGTGGCAACCAAGTCCCGG 1876

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:us-09-022-257-1

seq_documentation_block:
: Sequence 1, Application us/09022257
: Patent No. 6197525
: GENERAL INFORMATION:
: APPLICANT: Yao, Zhengbin
: APPLICANT: Spriggs, Melanie
: APPLICANT: Fanslow, William
: TITLE OF INVENTION: No. 6197525el Receptor That Binds IL-17
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: Apple Power Macintosh
: OPERATING SYSTEM: Apple Operating System 7.5.5
: SOFTWARE: Microsoft Word for Apple, Version 6.0.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/022,257
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/620,694
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/620,694
: FILING DATE:
: APPLICATION DATA:
: APPLICATION NUMBER: USSN 08/410,535
: FILING DATE: 23 MARCH 1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Perkins, Patricia Anne
: REGISTRATION NUMBER: 34,695
: REFERENCE/DOCKET NUMBER: 2617-B
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206)587-0430
: TELEFAX: (206)
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3288 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Mouse
: STRAIN: HVS13 receptor
: FEATURE:

```


NAME/REV: CDS
LOCATION: 121..2715
US-09-022-257-1

Alignment scores:
Quality: 174.50 Length: 582
Ratio: 0.687 Gaps: 28
Percent Similarity: 43.643 Percent Identity: 20.619

Alignment block:
US-09-899-471-2 x US-09-022-257-1 ..

Align seg 1/1 to: US-09-022-257-1 from: 1 to: 3288

```

224 LysAsnLeuThr.....GlyProGlnAsnIleThrLeuAsn..... 235
|||||
316 AAAAAGCTGACCCGCTCTCCCAAAAAACATCTATATCAATCTTAGTGT 365
.....HisThrAspLeuValProCysLeuCysIleGlnValT 248
366 TTCCTTACCCAGCAGCAGAGATATTAGTCCCTGTGTGCATGTTGAG..T 412
248 rpsSerLeuGluProAsp..... 253
413 GGACCTGTCAGACAGATGCCAGATCTCTGACTGAGGGTGCAGAGCTG 462
254 .....SerGluArg.....ValGluPheCysPr 261
463 TCCGCTCTGAGCTGAAACACCAATGAGCGCTGTGTCAAGTTC...CA 509
261 oPheArgGluAspProGlyAlaHisArgAsnLeuThrPheIleAlaArgL 278
|||
510 GTTCTGTCCATGCTGCAGCATCACCGTAACGGGTGGGTTTCTTCA 559
278 euArg...ValLeuSerProGlyValIleTrpGlnLeuAspAlaProCysCys 293
:::
560 GCCACTTTGTGGTAGATCTCTGGC..... 582
294 LeuProGlyLysValThrLeuCysTrpGlnAlaProAspGlnSerProCy 310
583 CAGGAGTATGAAAGTACTGT..... 603
310 sGlnProLeuValProProValProGlnLysAsn..AlaThrValAsnGlu 326
:::
604 .CACCACCTGCGGAGCCCATCTCTGATGGGACCCCAACCAATATCA 652
327 ProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuCysValGlnVa 343
:::
653 .....AGATCATCTTTGCTCCGTGACGTGT 675
343 lserThr..... 345
676 GAGGACAGCAGATGAGAGTACTACTCATGCGTGAAGCTCAGCAGAGCT 725
346 .....TrpGluLysValGlnLeuGlnAlaCysSer 355
726 TTGGGATCCCAACATCATCTGTGGAGACTTGGACACACAGCATCTGGCAG 775
356 TrpAlaAspSerLeuGlyProPheLysAspAspMetLeuLeuValGluMe 372
|||
776 TGG..... 778
372 LysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluPro..... 386
779 .....ACTTCACCTCTGTGAATGATCAATCCACCCCTTA 809
387 ..SerGlyCysThrProLeu...ProSerMetAlaSerThrArgAlaAla 401
:::
810 CCAGGCTCTGCTGAAAAGTTTCTCCGACTCAGAGACCAACAGCTGCTTGG 859
402 ArgLeuGlyGlnLeuLeuGlnAspPheArgSerHisGlnCysMetG1 418
:::

```

```

860 ATGTGCTTAAACAATATTTGCGCCAGCAGAAAGAATTCATCAGCGA 909
418 nLeuThrAsnAspAspAsnMetGlySerLeu..TrpAlaCys..... 431
|||
910 GCTAATGTCACATTCATCTTAAGCAAGTTTCACTGTGTGCTGCATCACA 959
431 ..... 431
960 CGTGCAGCTCCAGCCCTTCTTACAGACGCTGCTTAATGACTGTTAGAC 1009
431 ..... 431
1010 ACGCTGACTGTGCGCTGCCCAGTAATCTCAATACACAGTTCCCAAG 1059
432 PrometAspLysTrpIleHisArgArgTrpVal.....LeuValTrpLe 446
|||
1060 CCAGTTGCAGACTACATTT...CCCTGTGGGTATGAGCTCATCAGACT 1106
446 uAlaCysLeuLeuLeuAlaAlaLeuPhePheLeu..... 459
|||
1107 CATGCGCATTTCTGTGTGGATGTGTATCTGCTGATCATCTGTATGA 1156
460 .....LeuLeuLysLysAspArgArgLysAlaAlaArgLysSer... 472
1157 CCTGGAGGCTTTCTGGCGCCGATCAAGAGAAACATGCTGATGATCCAAA 1206
472 ..... 472
1207 ATCAATGGCATCTTCCCGTAGCAGACTGATCCCAACCCCTGAGGCC 1256
473 .ArgThrAlaLeuLeuLeuHisSerAlaAspGlyAlaGlyTrpGluArgL 489
1257 CAGAAAGCTGTGATGCTGTACTGTGCGCCGACCAACCCCTTATGTGAGG 1306
489 euValGlyAlaLeuAlaSerAlaLeu...SerGlnMetProLeuArgVal 504
:::
1307 TGGTCTAAAGTTCGCCAGATCTCTGATCATCTGCTGTCAGCTGAAGTA 1356
505 AlValAspLeuTrpSerArgArgGluLeuSerAlaHisGlyAlaLeuAl 521
|||
1357 GCCCTGACCTCTCGAAGACAGGTATCTCTGAGGTGGGGGTCAATGAC 1406
521 atPrPheHisGlnArgArgArgLLeuGlnGlnGlyValValI 538
|||
1407 CTGGGTGAGCCGACAGAGAGAGATGTGGAGAGCACTCAAAATCA 1456
538 lLeuLeuPheSer..... 542
1457 TCATCTGTGTCTCCGAGGCAACCCCAAGCAAGTGAAGACTATCTGGGT 1506
543 .....ProAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnLeuG1 556
1507 TGGGCTGAGCTGCTGTCCAGCTAGCGTGTGACCACTGA..... 1545
556 nThrValGluProGlyProHisAspAlaLeuAlaAlaTrpLeuSerCysV 573
1546 .....AAGCTGCT...GGGACCTTTTTCACCTGACGCAAGCAATGA 1585
573 alLeuProAspPheLeuGlnGlyArgAlaThrGlyArgValGlyVal 589
|||
1586 TCTGTCCAGACTTCAAGAGGCCAGCTCTCTGCGACCTACCTGTGTTTGC 1635
590 TyrPheAspGlyLeuLeuHisProAspSerValProSerProPheArgVa 606
|||
1636 TACTTCACTGATCTGTAGTAGAGGATGTCCCGACCTTCAACT 1685
606 lAlaProLeuPheSerLeuPro..ThrGlnLeuProAlaPheLeuAspAla 622
:::
1686 CACCTCCAGGTACCCACTCATGACAGATTGTGAGAGGTTTACTTCCGA 1735
623 LeuGln.....GlyLysCysSer.....Th 629
:::
1736 TCCAGAGCTCGAGATGTTTGAACCCCGGCGGATCCACCATCTCAGAGAG 1785

```



```

261 .....ProphetargLysAspProGlyAlaHisArg 270
784 AGATCTGCTGACCAAGTTTCCGACATGAGAACACAGTTGCTTGAG 833
271 AsnLeuTrpHisIleAlaArgLeuArgValLeuSerProGlyValTrpGlu 287
834 CACATGACACCAATA..... 848
287 nLeuAspAlaProCysCysLeuProGlyLysValThrLeuCysTrpGlnA 304
848 ..... 848
304 lAspAspGlnSerProCysGlnProLeuValProProValProGlnLys 320
849 .....CCTGCCGCCAGA... 860
321 AsnAlaThrValAsnGluProGlnAspPheGln..... 331
861 .....CCAGAGAGTTCCACAGCCAGTCACAGCTCAC 892
332 LeuValAlaGlyHisProAsnLeuCysValGlnValSerThrTrpGlnL 348
893 ACTCACCTACGACCACTTAAGGCTGCTGCG.....CACG 930
348 yValGlnLeuGlnAlaCysSerTrpAlaAspSerLeuGlyProPheLys 364
931 AAGTGACAGATCCAG.....CCCTTC... 950
365 AspAspMetLeuLeuValGlnMetLysThrLysAsnAsnTrpSerVal 381
951 .....TTGACGACGCTGCTCAATGAC..... 971
381 lCysAlaLeuGluProSerGlyCysThrProLeuProSerMetAlaSerT 398
972 .....TGCCTCAGACACTCCGCACTGTTCTGCCCCAGAAAGCCAGACA 1017
398 hArgAlaAlaArgLeuGlyGlnGluLeuGlnAspPheArgSerHis 414
1018 CT.....CCAGAACCAATTCGCGACTAC..... 1040
415 GlnCysMetClnLeuTrpAsnAspAspMetGlySerLeuTrpAlaCys 431
1041 .....ATGCCCTGTGG..... 1052
431 sPrometAspLysTrpIleHisArgArgTrpValLeuValTrpLeuAlaC 448
1053 .....GTGTACTGTTTCATCAGGCGCATCTCCA 1080
448 yLeuLeuLeuAlaAlaAlaLeuPhePheLeuLeuLysLysAsp 464
1081 TCCTGCGTGGGGCTCGCTCATC.....CTGCTCATGCTGCGATGACC 1124
465 ArgArgLysAlaAlaArgGlySer..... 472
1125 TGGAGGCTACTGCTGGAAAGTGAATAATACAGTATGACACCAATA 1174
473 .....A 473
1175 CACCGATGCGCTCGCTCGCTGACCTGATCCCCACCGCTGAAGCCA 1224
473 rghrAlaLeuLeuHisSerAlaAspGlyAlaGlyTrpGluArgLeu 489
1225 GGAAGGCTGATCATCTACTACACCCAGCCCTCTAGCTGACGCG 1274
490 ValGlyAlaLeuAlaSer...AlaLeuSerGlnMetProLeuArgValAl 505
1275 GTCTGAAATTCGCCAGTTCTCTCATCAGCTGCGGCGACGGAAGTGGC 1324
505 aValAspLeuTrpSerArgArgGluLeuSerAlaHisGlyAlaLeuAlaT 522
1325 CCTGACCTCTGAGAGAGCGGCTCTCGAGGACAGAGTATCATGACCT 1374

```

```

522 rPheHisGlnArgArgGlyLeuGlnGlnGlyValIle 538
1375 GGGTGGCGCTCAGAGAGAGATGGTGAGAGCAACTATGATCATC 1424
539 LeuLeuPheSerProAlaAlaValAlaGlnCysGlnGlnTrpLeu..... 553
1425 GTCTGTCTCCCGCGGCGGCGCAAGTGAGCAGCGCTCTGGGCGC 1474
554 .....GlnLeuGlnThrValGluProGlyPro...HisAspA 565
1475 GGGGCGCGCTGCGGCTGCTGCGACCAAGAACCCGCGGAGCC 1524
565 lAlaAlaAlaTrpLeuSerCysValLeuProAspPheLeuGlnGlyArg 581
1525 TGTTCACAGCAACCATGAACATGATCTCCGCACTTCAAGAGCCAGCC 1574
582 AlArtrGlyArgTrpValGlyValTrpPheAspGlyLeuLeuHisProAs 598
1575 TGCTTCGCGACCTACGTAGTGTCTACTTACAGGAGCTGAGTGCAGG 1624
598 pSerValProSerProPheArgValAlaProLeuPheSerLeuProThrG 615
1625 CGAGCTCCCGCACTGTTGCGCGCGCGCGGTAACCCGCTC..... 1667
615 lLeuProAlaPheLeuAspAlaLeuGlnGlyLysSerThrSerAla 631
1667 ..... 1667
632 GlyArgProAlaAspArgValGluArgValThrGlnAlaLeuArgSerA 648
1668 .....ATGACACAGTTCGAGAGAGTGTACTTCCGATCCAGACCT 1708
648 lAlaLeuSerCysThrSerSerGlnAlaProGlyCys 662
1709 GGAGTGTTCACAGCGCGGCGGCGCATGACCGCGTAGGGGAGCTGT 1752

seq_name: /cgn2_6/protdata/2/ina/6A_COMB.seq:US-09-022-255-9
seq_documentation_block:
; Sequence 9, Application US/09022255
; Patent No. 6072033
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022.255
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: USN 08/620,694
; FILING DATE: 21 MARCH 1996
; APPLICATION NUMBER: USN 08/538,765
; FILING DATE: 7 AUGUST 1995
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: USN 08/410,535
; FILING DATE: 23 MARCH 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne

```


480 GAGTTTCTGTCCAACTGAGGCATCACCACAGGCGGTGGCGTTTACCTT 529

[illegible]


```

1053 ..... 111 .....
448 yslleuLeuLalaalaleuPhePheLeuLysLysasp 464
      ||||| .....
1081 TCGCTGCTGGGCTCCGCATC..... CTGCTCATCTGTCATGAC 1124
465 ArgArgyslaalAArgLysr..... 472
      ||| |||||
1125 TGGAGCTAGCTGGGCTGGAAGTGAATAATACGTGATGACCAATA 1174
473 ..... A 473
1175 CACGATGGCTGCTGGGCTGACCTGATCCGCCACCGCTGAAGCCA 1224
473 ryrhAlaLeuLeuHisSerAlaAspGlyAlaGlyTyrGlyArgLeu 489
      || ::::: ||||| |||
1225 GGAAGGTCTGATCATCTACACCGCAGCACCCCTCATCGTGAAGTG 1274
490 ValGlyAlaLeuAlaSer..AlaLeuSerGlnMetProLeuArgValAl 505
      ||||| ::::: |||||
1275 GTCTGTAAATTCGCCCAATCTCTGCTACACGCTCGCGCAGCAAGTGGC 1324
505 aValAspLeuTyrSerArgArgGluLeuSerAlaHisGlyAlaLeuAlaT 522
      ||||| ::::: |||||
1325 CCTGGACCTGCTGGAAGCAAGCCATCTCGGAGCAGGAGTATGACCT 1374
522 rPpHeHisGlnArgArgArgIleGlnGlnGlyGlyValAlaIle 538
      || ::::: ||||| |||||
1375 GGGGAGGCGCCAGAGAAGAGTGTGAGAGCAACTTAAGATATC 1424
539 LeuLeuPheSerProAlaAlaValAlaGlnCysGlnGlnTyrLeu.... 553
      ::::: ||||| ::::: |||||
1425 GTCCTGTGCTCCCGCAGCGCGCCAGGTGACGCGCTCTGGGCGC 1474
554 ..... GlnLeuGlnThrValGluProGlyPro..HisAspA 565
      ||||| ::::: |||||
1475 GGGGCGCCTGTGGGCTGGCTGGCAGCACAGMAAGCCCGTGGGAGCC 1524
565 lAlaAlaAlaTyrLeuSerCysValLeuProAspPheLeuGlnGlyArg 581
      ::::: ||||| ::::: |||||
1525 TGTTCACCTGCGACCATGAACATGATCTCCCGACTTCAGAGGCCAGCC 1574
582 AlaThrGlyArgTyrValGlyValTyrPheAspGlyLeuLeuHisProAs 598
      ::::: ||||| ||||| :::::
1575 TGCCTGCGACCTACGTAGTGTGCTGCTACCTCACCGAGGTCAAGCTGACGG 1624
598 pSerValProSerProPheArgValAlaProLeuPheSerLeuProThrG 615
      ::::: ||||| ::::: |||||
1625 CGAGCTCCCGACCTGTTCGGCGCGCGCGCGGTACCCGCTC..... 1667
615 lIneuProAlaPheLeuAspAlaLeuGlnGlyGlyCysSerThrSerAla 631
1667 ..... 1667
632 GlyArgProAlaAspArgValGlu..ArgValThrGlnAlaLeuArgSerA 648
      ||||| ||| ||| ||||| |||||
1668 ..... ATGCAACAGTTTGGAGAGGTGTACTTCCGATCTCCAGGACCT 1708
648 lAlaLeuAspSerCysThrSerSerSerGlnAlaProGlyCysCys 662
      ||| ::::: ||| ||| |||
1709 GGAAGATGTTCCACGCGCGCGCATCACCGGCTAAGGGAGAGCTGT 1752
name: /cgn2_6/ptodeta/2/lna/dA_COMB_seq:US-09-022-696-9

```

seq. documentation block:
 ? Sequence 9, Application US/09022696
 ? Patent No. 6072037
 ?
 ? GENERAL INFORMATION:
 ?
 ? APPLICANT: Yao, Zhenqin
 ? APPLICANT: Spriggs, Melanie
 ? APPLICANT: Fanirow, William
 ?
 ? TITLE OF INVENTION: No. 6072037-e1 Receptor That Blinds IL-17

```

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-09-022-696
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/620,694
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEX: (206)
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 3223 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Human
STRAIN: IL-17 R (MCTLA8 receptor)
FEATURE:
NAME/KEY: CDS
LOCATION: 93..2693
US-09-022-696-9

alignment_scores:
Quality: 154.00 Length: 732 Gaps: 33
Ratio: 0.552 Percent Identity: 19.399
Percent Similarity: 38.115

alignment_block:
US-09-899-471-2 x US-09-022-696-9 ..
Align seg 1/1 to: US-09-022-696-9 from: 1 to: 3223

55 SerAlaProGlyPro.....Valle 61
   ::::| | | | |
120 GGCGTCCGGGGGCCCTGCCTGTCTTCCTCCGTGGCGCTGCT 169
    :      |||||
61 uVAlPrO.....ThArgLeuGlThrGIueValLeuArgc 74
   |:|||||         |||||
170 GGGCCCCGGTGCGCCCTCCGCACTCTGGACACAGCGGCGCTGTCT 219
     |::::|              |||
74 ySfRoglInIstThrAsPcySaLaLeuArGvaLargVaIvalAlHisLeu 90
   ||:|||             :::|:|:|:|
220 GCACCAGCG...GGCTAAACTGCAAGGTCAAAGAATATAACC 260
     |||               |||
91 AlavAlHiselYlIsTPPaLIuProClugluValaglYLysSeraspse 1077
```



```

261 ..:::.. || ..:::.. 290
261 TCCCTGGATGACGCTGATTCACCTCGA.....
107 rcIuLeuIngluSerArgAsnIaSerIeuInlaIgluValIleus 124
107 ..:::.. || ..:::..
291 .AACCTGACCCCTCTCCCAAGACCTGACAGATCCAGTGCACCTTG 339
124 erPhe.....GlnAlaTyrProIleAlaArgCysAlaLeu 136
124 ..:::.. || ..:::..
340 CCCACACCCACAGAGAGACCTGTCCTCCCGCTGCAC.....ATC 380
137 GluValGlnValProAlaAspIleuValGlnProGlyGlnSerValGlyse 153
137 ..:::.. || ..:::..
381 GAATGGACACTGACAGACAGC..... 401
153 rAlaValPheAspCysPheGlnIaSerIeu.....GlyAlaG 166
153 ..:::.. || ..:::..
402 .....GCCAGCATCTGTACCTCGAGGCTGCAG 429
166 luValGlnIle..... 169
166 ..:::.. || ..:::..
430 AGTATCTCTCTCGACGCTGACACCAATGACCTTGTGCGTACGTTT 479
170 ..... 173
170 ..:::.. || ..:::..
480 GAGTTTCTGTCCAACTGAGCATCACACAGGGGGTGGCTTACTCT 529
174 .....LysProArgTyrGlnIleu.....A 182
174 ..:::.. || ..:::..
530 CAGCCACTTGTGTGTTGACCTGACCCAGAAATGAGTGACCTGCAC 579
182 snLeuThrGlnIleuProAspGlyAspAsnValIleuLeuThrLeuAsp 198
182 ..:::.. || ..:::..
580 ACCTGCCACGCCATCCTGATGGGAGC..... 608
199 ValSerGluGlnAspPheSerPheLeuLeuTyrLeuArgProValPr 215
199 ..:::.. || ..:::..
609 ...CCAAACACACAGTCCAAATTCCTT.....GTGCC 640
215 oASP.....AlaLeuL 219
215 ..:::.. || ..:::..
641 TGACGTGTGACAGCCAGATGAGTAAACACGCCATGCATGATGACAG 690
219 ySerLeuTyrTyrIleuAsnLeuThrGlyProIleuAsnIleThrLeuAsn 235
219 ..:::.. || ..:::..
691 GCAGCTGTGGACCCCAACATCACC..... 716
236 HisThrAspLeuValProCysLeuCysIleGlnValTyrSerLeuGluPr 252
236 ..:::.. || ..:::..
717 .....GTGGAGACCTGGAGGC 733
252 oAspSerGluArgValGluPheCys..... 260
252 ..:::.. || ..:::..
734 CCACAGGTGCTGTGAGCTTCACCTGTGAAGACATCAACCATTAAC 783
261 .....ProPheArgGluAspProGlyAlaHisArg 270
261 ..:::.. || ..:::..
784 AGATCTGCTGACCACTTTTCCGACATGAGAACACACAGTTCCTTGA 833
271 AsnLeuTyrPheHisIleAlaArgLeuArgValIleuSerProGlyValTyrGln 287
271 ..:::.. || ..:::..
834 CACATGACACACATA..... 848
287 nLeuAspAlaProCysCysLeuProGlyValIleuLeuCysTyrGlnA 304
287 ..:::.. || ..:::..
848 ..... 848
304 laProAspGlnSerProCysGlnProLeuValProProValProGlnLys 320
849 .....CTGGGCCACAGA... 860
321 AsnAlaThrValAsnGluProGlnAspPheGln..... 331
321 ..:::.. || ..:::..
861 .....CCAGAAAGTTCACACAGGATCCAACTGAC 892
861 ..:::.. || ..:::..
332 .LeuValAlaIleGlnHisProAsnLeuCysValGlnValSerThrProGluL 348
332 ..:::.. || ..:::..
893 ACTGACTTAACGACCACTTAAGGCTGCTGCC.....CACG 930
348 yValGlnIleuGlnAlaCysSerThrAlaAspSerLeuGlyProPheLys 364
348 ..:::.. || ..:::..
931 AACTGCAGATCCAG.....CCCTTC... 950
365 AspAspMetLeuLeuValGluMetLysThrGlyLeuAsnAsnThrSerVa 381
365 ..:::.. || ..:::..
951 .....TTCAGCAGCTGCCATATGAC..... 971
381 LysAlaLeuGluProSerGlyCysThrProLeuProSerMetAlaSerT 398
381 ..:::.. || ..:::..
972 ....TCCCTCAGACACTCCGCACTTTCCTGCCCAAAATGCACAGACA 1017
398 hrArgAlaAlaArgLeuGlyGluGluLeuLeuGlnAspPheArgSerHis 414
398 ..:::.. || ..:::..
1018 CT.....CCAGAACCAATTCGGACTAC..... 1040
415 GlnCysMetGlnLeuTyrPheAsnAspAspAspMetGlySerLeuTyrAlaCy 431
415 ..:::.. || ..:::..
1041 .....ATGCCCTGTGG..... 1052
431 sProMetAspLysTyrIleHisArgArgTyrValLeuValTyrPheAlaC 448
1053 .....GTGTACTGTGTATCATCAGGCGCATCTCCA 1080
448 yLeuLeuLeuAlaAlaAlaLeuPhePheLeuLeuLysAsp 464
448 ..:::.. || ..:::..
1081 TCCCTGCTGGTGGCTCCGTCATC.....CTGTCAATGTCATGCATGACC 1124
465 ArgArgLysAlaAlaArgLysSer..... 472
465 ..:::.. || ..:::..
1125 TGGAGGCTGAGCTGGCTGGAGTGAATAATACATGATGACACCAATA 1174
473 .....A 473
1175 CACCGATGGCTGCTGCGGCTGACCTGATCCCGCCACCGCTGAAGCCA 1224
473 rGThrAlaLeuLeuLeuHisSerAlaAspGlyAlaGlyTyrGluArgLeu 489
473 ..:::.. || ..:::..
1225 GGAAGTCTGATCATCTACTACACCGCACCCCTCTGATGAGAGCTG 1274
490 ValGlyAlaLeuAlaSer...AlaLeuSerGlnMetProLeuArgValAl 505
490 ..:::.. || ..:::..
1275 GTCTGGAATTCGCCCAATTCCTGCTACCGCTGCGGACAGGAGTGGC 1324
505 aValAspLeuTyrSerArgArgGluLeuSerAlaHisGlyAlaLeuAlaT 522
505 ..:::.. || ..:::..
1325 CCTGGAGCTGCTGGAGAGAGAGCATCTCGAGGCGAGAGTGTAGACT 1374
522 rPheHisIleGlnArgArgArgIleLeuGlnGluGlyGlyValValIle 538
522 ..:::.. || ..:::..
1375 GGGTGGCGGCTGAGAACAGAGATGCTGGAGAGCACTCTGATCATC 1424
539 LeuLeuPheSerProAlaAlaValAlaGlnCysGlnIleThrLeu... 553
539 ..:::.. || ..:::..
1425 GTCTGTGCTCCCGGCGGACAGCGGCCAAGTGGACGCTCTGGGCGG 1474
554 .....GlnLeuGlnThrValGluProGlyPro...HisAspA 565
554 ..:::.. || ..:::..
1475 GGGGCGGCTGTGGCGCTGCGGTGGACACAGGAAAGCCGTGGGGGACC 1524
565 laLeuAlaAlaThrPheLeuSerCysValLeuProAspPheLeuGlnGlyArg 581
565 ..:::.. || ..:::..
1525 TGTTCATGTCACCAATGACATGATCTCCCGGATTCAGAGAGGCGGACC 1574
582 AlaThrGlyArgTyrValGlyValTyrPheAspGlyLeuLeuHisProAs 598
582 ..:::.. || ..:::..
1575 TGTCTGGCACTACGTAGTGTCTGCTACTTCAGCGAGGTCACTGTGACGG 1624

```



```

alignment_block:
US-09-899-471-2 x US-08-978-773-3  ..
Percent Similarity: 38.115    Percent Identity: 19.399

Quality: 154.00    Length: 732
Ratio: 0.552    Gaps: 33
Percent Similarity: 38.115    Percent Identity: 19.399

Align seg 1/1  to: US-08-978-773-3  from: 1  to: 3223

55 serialProGlyPro.....Valle 61
   :::::|||||
120 GCGTCCCGGGGCCCTGCTGCTCCTCGCTCTCGGCGCGCT 168
   :::::|||||

61 valPro.....ThraGlauGlnThrGluValLeuArgc 74
   |:::::|||||
170 GGGCGGGGGCGGCCCTCCCTCGCGCTGTGACACCGGCGCTGT 218
   |:::::|||||

74 ysdProGlnYthrAspCysAlaLeuArgValArgValValAlaHisLeu 90

```

```

Align seg 1/1 to: US-08-978-773-3 from: 1 to: 3223

55 serialProglyPro.....Valle 61
   :::::|||||
120 GCGTCCCGGGGGCCCTGGGGCTGTCTCTGCTCTCTGGGGCGTCT 169
61 uValPro.....ThArgLeuGlnThrGluValLeuAArgC 74
   :::::|||||
170 GGGCGGGGGGGGGGGCCCTCCCTGGAGCTCTGGACCAACGGGGCTGTCT 219
74 ysbProGlnLysThrAspCysAlaLeuAArgValAArgValValHisLeu 90
   :::::|||||
220 GCTCCACAGCGG.....GGGCTAAATCGACAGGTCAAGATAGTACC 260
91 AlAValHisGlnHisThrAlaGluProGlnGluAlaGluLysSerAspse 107
   :::::|||||
261 TGGCTGGATGACAGCTGGATTCACTCTGA..... 290
107 rgluLeuGlnGluSerAArgAsnAlaSerLeuGlnAlaGluValLeuAs 124
   :::::|||||
291 AACCTGACCCCTCTCCCAAGAGAGCTGTCCCGGTGCTAC.....ATC 380
124 eArPhe.....GlnAlaLysProIleAlaArgCysAlaLeu 136
   :::::|||||
340 CCCACACCCACAAGAGAGAGCTGTCCCGGTGCTAC.....ATC 380
137 GluValGlnValProAlaAspLeuValGlnProGlnGlnSerValGlyse 155
   :::::|||||
381 GAATGACACTGACAGACAGAC..... 401
153 rAlAValPheAspCysPheGlnAlaSerLeu.....GlyAlaG 166
   :::::|||||
402 .....GCCAGATCTCTGACTCGAGGGGTGAC 429
166 lValAlaGlnLe..... 169
   :::::|||||
430 AGTTACTGTCTGACACTGTAACACCAATGAAGCTTGTGCTCAGGTTT 479
170 .....TTPSerTyrThr 173
480 GAGTTCTGTCCAAATGAGGACATACCACAGCGGTGGCGTTTACCTT 529
174 .....LysProAArgTyrGlnLysGluLeu.....A 182
530 CAGCCACTTGTGTGGTACCGCTGACAGAGATATGAGGTGACCGTTAC 579
182 snLeuThrGlnGlnLeuProAspCotLysAsnValLeuLeuThrLeuAsp 198
   :::::|||||
580 ACCTGGCCAAGCCCATCTCTGATGGGGAC..... 608
199 ValSerGlnGlnGlnAspPheSerPheLeuLeuTyrLeuAArgProValPr 215
   :::::|||||
609 .....CCAAACCAACAGTCCAAAGATTCTT.....GTGCC 640
215 oAsp.....AlaLeuL 219
   :::::|||||
641 TCACCTGTGACGACGCGCAGATGAAGTGAACACGCGCATGACGACTCAG 690
219 ysserLeuTyrTyrLysAsnLeuThrGlyProGlnAsnIleThrLeuAsn 235
   :::::|||||
691 GCAGCGTGTGGACCCCAACATCACC..... 716
236 HlSThrAspLeuValProCysLeuCysIleGlnValITPserLeuGluPr 252
   :::::|||||
717 .....GTGGAGACCTGTGAGC 733

```



```

252 oAspSerGluArgValGluPheCys..... 260
734 CCACAGCTGCTGTGAGCTTCACCTGTGAAACAATTCACCATACC 783
261 ..... 270
784 AGATCCGCTGACCAAGTTTCCGACATGAGAACACAGTTGCTTGGAG 833
271 AsnLeuThrPheIleAlaArgLeuArgValLeuSerProGlyValTyrPcl 287
834 CACATGACACCAATA..... 848
287 nLeuAspAlaProCysCysLeuProGlyLysValThrLeuCysTyrPcl 304
848 ..... 848
304 IaProAspGlnSerProCysGlnProLeuValProProValProGlnLys 320
849 ..... 860
321 AsnAlaThrValAsnGluProGlnAspPheGln..... 331
861 ..... 892
332 LeuValAlaGlyHisProAsnLeuCysValGlnValSerThrTyrPcl 348
893 ACTCACTCAAGCACTTAAGGCTGCTGCC..... CACC 930
348 ystValGlnLeuGlnAlaCysSerThrAlaAspSerLeuGlyProPheLys 364
931 AAGTCGAGATCCAG..... CCGTTC... 950
365 AspAspMetLeuLeuValGluMetLysThrGlyLeuAsnAsnThrSerVa 381
951 ..... 971
381 LysAlaLeuGluProSerGlyCysThrProLeuProSerMetAlaSerT 398
972 ..... 1017
398 htrArgAlaAlaArgLeuGlyGluGlnLeuGlnAspPheArgSerHis 414
1018 CT..... CCAGAACCAATTCGGACTAC..... 1040
415 GlnCysMetGlnLeuTrpAsnAspAspAsnMetGlySerLeuTyrPcl 431
1041 ..... 1052
431 sPrometAspLysTyrIleHisArgArgTyrValLeuValTyrPcl 448
1053 ..... 1080
448 ystLeuLeuLeuAlaAlaAlaLeuPhePheLeuLeuLysLysAsp 464
1081 TCCTGCTGGGCTCGCATC..... CTGCTCATGCTCTGCATGACC 1124
465 ArgArgLysAlaAlaArgGlySer..... 472
1125 TGAAGCTACTGCGCTGGAAGTGAATAATACAGTATGACACCAATA 1174
473 ..... 473
1175 CACCGATGCGCTGCGCTGAGCTGATCCGCCACCGCTGAAAGCCA 1224
473 rghrAlaLeuLeuHisSerAlaAspGlyAlaGlyTyrGlnArgLeu 489
1225 GGAAGCTGAGATATCTACTACACCAACCCCTTACGTGAGACTG 1274
490 valGlyAlaLeuAlaSer...AlaLeuSerGlnMetProLeuArgValAl 505
1275 GTCCTGAATTCGCCAGTTCCTGCTACCGCCCGCAGCAAGTATGC 1324
505 aValAspLeuTyrSerTrpArgGlnLeuSerAlaHisGlyAlaLeuAlaT 522

```

```

1325 CCTGACCTGCTGGAAGACAGCAGCATCTCGGAGCAGAGTATACCT 1374
522 rPheHisGlnAlaArgArgIleLeuGlnGluGlyValValIle 538
1375 GGTGGGCGCTGAGAACGAGAGATGTTGAGAGCAACTTAAGATCATC 1424
539 LeuLeuPheSerProAlaAlaValAlaGlnCysGlnGlnTyrPleu 553
1425 GTCTGCTCTCCGCGGACGCGCGCAAGTGGACGCGCTCTGGGCGC 1474
554 ..... GlnLeuGlnThrValGluProGlyPro...HisAsp 565
1475 GGGGCGCGCTGCGCGCTGCGCTGCGACCAAGAACCCGCTGGGAGCC 1524
565 lAlaAlaAlaTyrPleuSerCysValLeuProAspPheLeuGlnArg 581
1525 TGTTCACCTGACGACCAAGAACATGATCTCCCGACTTCAAGAGCCAGCC 1574
582 AlaThrGlyArgTyrValGlyValTyrPheAspGlyLeuLeuHisProAs 598
1575 TGTTCGCGACCTACGTAGTCTGCTACCTCAGCGAGGTGAGTGCAGCG 1624
598 pSerValProSerProPheArgValAlaProLeuPheSerLeuProThc 615
1625 CGAGTCTCCGACCTGTTGCGCGCGCGGTACCGCTC..... 1667
615 lLeuProAlaPheLeuAspAlaLeuGlnGlyCysSerThrSerAla 631
1667 ..... 1667
632 GlyArgProAlaAspArgValGluArgValThrGlnAlaLeuArgSera 648
1668 ..... ATGACAGGTTCCAGAGAGGTGTACTTCCGATCAGAGACT 1708
648 lAlaLeuSerCysThrSerSerSerGluAlaProGlyCysCys 662
1709 GGAATGTTCAGCGCGCGCATGACCGCTAGGGAGCTGT 1752

seq_name: /cgn2_6/plodata/2/lna/6A_COMB.seq:US-09-022-253-9
seq_documentation_block:
; Sequence 9, Application US/09022253
; Patent No. 6096305
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 6096305el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,253
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,694
; FILING DATE: 21-MARCH-1996
; APPLICATION NUMBER: USSN 08/538,765
; FILING DATE: 7 AUGUST 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```


APPLICATION NUMBER: USSN 08/410,535

FILING DATE: 23 MARCH 1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,695

REFERENCE/DOCKET NUMBER: 2617-B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)587-0430

TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 323 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Human

FEATURE: IL-17 R (IL17A receptor)

NAME/KEY: CDS

LOCATION: 93..2693

US-09-022-253-9

alignment_scores:

Quality: 154.00

Ratio: 0.552

Percent Similarity: 38.115

Percent Identity: 19.399

Length: 732

Gaps: 33

Align seg 1/1 to: US-09-022-253-9 from: 1 to: 3223

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

US-09-899-471-2 x US-09-022-253-9

170TTPSerThr... 173
 480 GAGTTCTGTCCAAAGTGGATCAGACAGGCGGTGGCTTACCTT 529
 174LysProArgTyrGlnGlu... 182
 530 CAGCAGCTTTGGTGGTACCTGACAGGATATGAGTGACCTTACC 579
 182 snLeuThrGlnGluLeuProAspGlyAspAsnValLeuThrLeuasp 198
 580 ACCTGCCCAAGCCCAATCCCTGATGGGAC... 608
 199 ValSerGluGlnGlnAspSerPheLeuLeuThrLeuArgProValPr 215
 609 ...CCAAACACAGTCCAAAGATTCCTT...GTGCC 640
 215 OASP...AlaLeuL 219
 641 TCACCTGTGACGACGCGCAGATGAGTAACACGCCATGATGAGCTCAG 690
 219 ysSerLeuTyrTyrLysAsnLeuThrGlyProGlnAsnIleThrLeuasn 235
 691 GCAGCCTGTGGAGCCCAACATCACC... 716
 236 HisThrAspLeuValProCysLeuCysIleGlnValTyrSerLeuGluPr 252
 717GTGGAGCCTGGAGGC 733
 252 OASPserGluArgValGluPheCys... 260
 734 CCACAGCTGCTGTGAGCTTACCTGTGAGAACGATACCATTCATACC 783
 261ProPheArgGluAspProGlnAlaHisArg 270
 784 AGATCCTGCTGACAGATTTCGACATGAGAGAACCAACCATGCTTGGAG 833
 271 AsnLeuThrPheIleAlaArgLeuArgValLeuSerProGlyValTyrGln 287
 834 CACATGACACACATA... 848
 287 nLeuAspIleProCysCysLeuProGlyLysValThrLeuCysTyrPola 304
 848 848
 304 IAProAspIleSerProCysGlnProLeuValProProValProGlnLys 320
 849CCTGGCCAGCA... 860
 321 AsnAlaThrValAsnGluProGlnAspPheGln... 331
 861CCAGAGAGTTCACACAGATCCAACGTCCAC 892
 332 LeuValAlaGlnHisProAsnLeuCysValGlnValSerThrTyrPola 348
 893 ACTACCTGTACGCAACCTTAAAGGTGCTGTCC...CACC 930
 348 ysValGlnLeuGlnAlaCysSerTyrPalaAspSerGluGlyProPheLys 364
 931 AAGTCCAGATCCAG...CCCTTC... 950
 365 AspAspMetLeuLeuValGlnMetLysThrGlyLeuAsnAsnThrSerVa 381
 951TTCAGACCTGCTCAATGAC... 971
 381 lCysAlaLeuGluProSerGlyCysThrProLeuProSerMetAlaSerT 398
 972 ...TGCCTCAGACACTCCGAGCTGTTCCTGCCAGAAATGCCAGACA 1017
 398 hrArgAlaAlaArgLeuGluGluLeuLeuGlnAspPheArgSerHis 414
 1018 CT...CCAGAACCAATTCGGAGCTAC... 1040


```

GENERAL INFORMATION:
APPLICANT: Yao, Zhengbin
APPLICANT: Spriggs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6100235el Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,260
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/620,694
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USN 08/410,535
FILING DATE: 23 MARCH 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 3223 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Human
STRAIN: IL-17 R (hCTLA8 receptor)
FEATURE:
NAME/KEY: CDS
LOCATION: 93..2693
US-09-022-260-9

alignment_scores:
Quality: 154.00 Length: 732
Ratio: 0.552 Gaps: 33
Percent Similarity: 38.115 Percent Identity: 19.399

alignment_block:
US-09-899-471-2 x US-09-022-260-9 ..

Align seg 1/1 to: US-09-022-260-9 from: 1 to: 3223

55 Serialaprotglypro.....Valle 61
:::|||||
120 GTGTGTCGGGGGCCCTGCTGGGCTGCTCTGCTGCTCGGCGCTGCT 169
61 ValPro.....ThArgLeuGIaThrGIuLeuValIeuAarg 74
||||| ::|||
170 GGCCCGGGGTGGCGCTCCTCGACACTCTTGAGACCACGGGGCGCTGCTT 219

```


1525 TGTTCACCTGCAGCCATGACATGATCCTCCCGACTTCAAGAGCCAGCC 1574
582 AlaThrGlyArgTyrValGlyValTyrPheAspGlyLeuLeuHisProAs 598
::: ||| ||||| |||||::: |||||:::
1575 TGCCTCGGCACCTACCTAGTCTCTACTTCAAGAGGTCACTGTGACGG 1624
598 pSerValProSerProPheArgValAlaProLeuPheSerLeuProThrg 615
:::|||||::: ||| :::||||| :::
1625 CGACGTCCCGCACCTGTTCGGCGCGCGCGGTACCCGCTC..... 1667
615 InLeuProAlaPheLeuAspAlaLeuGlnGlyLysSerThrSerAla 631
1667 1667
632 GlyArgProAlaAspArgValGlu ArgValThrGlnAlaLeuArgSera 648
||| ||| ||| |||::: ||| |||:::
1668ATGGACAGGTTCCAGAGAGTGTCTTCCGCAATCCAGGACCT 1708
648 lalaAspSerCysThrSerSerSerGluAlaProGlyCysCys 662
||| :::: ||| ||| |||
1709 GGAGATGTTCCAGCGCGCGCGCATGCACCGCGTACGGGAGCTGT 1752

of: US-09-899-471-2 to: EST: * out_format : pfs

Date: Sep 28, 2002 9:43 PM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framed_p2n.model -DEV=xlh
-O=/cgr2_1/USPFO.spool/US09899471/runat.27092002.142439_17896/app_query.fasta.1.2519
-DB=EST -OPMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blomsum62 -TRANS=human4.0.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0
-ALIGN=15 -MODE=LOCAL -OUTPR=pfs NORM=ext -HEAPSIZ=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09899471.ecgml.1.7278
-NCPU=6 -ICPU=3 -LONGLOG -DEV TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPEX -WAIT -THREADS=1

Search information block:

Query: US-09-899-471-2

Query length: 674

Database sequences: 13736207

Database length: -1841457050

Search time (sec): 7183.320000

Score list:

Sequence	Strd Orig	ZScore	EScore	Len	Document
901.est1:AV010326	+ 1418.50	2361.61	2.7e-122	1039	AV010326 AV010326 Mus musculus
902.est1:BI1332055	+ 1226.00	2042.40	1.6e-104	697	BI1332055 602982878F1 NCI CGAP
903.est1:BG968034	+ 1187.00	1975.54	8.6e-101	784	BG968034 602832412F1 NCI CGAP
904.est1:BB609618	+ 1145.50	1903.01	9.4e-97	1006	BB609618 BB609618 RIKEN full-1
905.est1:AI0007139	+ 1095.00	1821.00	3.5e-92	775	AI0007139 us81f02.1 Soares mam
906.est1:BF237179	+ 1058.00	1759.92	9.2e-89	722	BF237179 BF237179 RIKEN full-1
907.est1:BB65706	+ 1039.00	1728.92	4.7e-87	722	BB65706 BB65706 RIKEN full-1
908.est1:BB637050	+ 1007.00	1675.53	4.4e-84	643	BB637050 BB637050 RIKEN full-1
909.est1:BE290395	+ 999.00	1623.16	2.2e-83	562	BE290395 BB637050 RIKEN full-1
910.est1:BF160122	+ 975.00	1621.47	4.5e-81	635	BF160122 601089427F1 NCI CGAP
911.est1:BI251117	+ 967.50	1605.78	3.4e-80	839	BI251117 602994315F1 NCI CGAP
912.est1:BG761197	+ 912.50	1512.11	5.6e-75	935	BG761197 602718928F1 NIH_MGC_45
913.est1:BI282622	+ 893.00	1483.89	2.1e-73	619	BI282622 UI-R-CW05-cdd-b-11-0-0
914.est1:BM210218	+ 876.00	1456.40	7.1e-72	561	BM210218 C0660H10-3 NIA mouse
915.est1:BM009552	+ 872.50	1448.04	2.1e-71	702	BM009552 603650118F1 NIH_MGC_41
916.est1:BI765330	+ 866.50	1434.94	1.1e-70	922	BI765330 603050641F1 NCI CGAP
917.est1:BM010292	+ 838.00	1387.56	4.8e-68	878	BM010292 603631169F1 NIH_MGC_41
918.est1:BG917539	+ 826.50	1369.44	5.0e-67	787	BG917539 602817086F1 NCI CGAP
919.est1:BB637050	+ 824.00	1368.07	5.3e-67	609	BB637050 2M0281A07R mouse 10kb
920.est1:BI689749	+ 821.50	1360.46	1.6e-66	829	BI689749 603316255F1 NCI CGAP
921.est1:BI689749	+ 805.50	1336.20	3.5e-65	653	BI689749 601767930F1 NCI CGAP
922.est1:W13694	+ 791.00	1305.88	1.7e-63	1118	W13694 ms94902.r1 Soares mouse
923.est1:BI770219	+ 785.00	1298.76	4.3e-63	855	BI770219 603053287F1 NIH_MGC_12
924.est1:BB760916	+ 778.00	1292.99	9.0e-63	497	BB760916 BB760916 RIKEN full-1
925.est1:BI820026	+ 771.50	1273.67	8.3e-62	886	BI820026 603037207F1 NIH_MGC_11
926.est1:BI812635	+ 754.50	1248.27	2.8e-60	736	BI812635 603082012F1 NIH_MGC_12
927.est1:BG867616	+ 754.00	1248.33	2.8e-60	794	BG867616 602786933F1 NCI CGAP
928.est1:BI913618	+ 753.00	1244.64	4.4e-60	880	BI913618 603180205F1 NIH_MGC_12
929.est1:BI913618	+ 751.00	1246.97	3.3e-60	526	BI913618 602986759F1 NCI CGAP
930.est1:AI463877	+ 748.00	1244.36	4.6e-60	422	AI463877 va32a03.y1 Guaywood
931.est1:BI770842	+ 727.50	1201.61	1.1e-57	893	BI770842 603058331F1 NIH_MGC_12
932.est1:BI767178	+ 726.00	1199.69	1.4e-57	846	BI767178 603060544F1 NIH_MGC_11
933.est1:BB736394	+ 724.00	1203.45	8.7e-58	444	BB736394 BB736394 RIKEN full-1
934.est1:BI9490389	+ 724.00	1196.81	2.0e-57	810	BI9490389 603031886F1 NIH_MGC_11
935.est1:BI919336	+ 719.00	1187.80	6.5e-57	855	BI919336 603032352F1 NIH_MGC_11
936.est1:BI919336	+ 716.50	1183.51	1.1e-56	862	BI919336 603037722F1 NIH_MGC_11
937.est1:AV752013	+ 713.50	1178.71	2.1e-56	843	AV752013 AV752013 NPD Homo sapi
938.est1:AM045329	+ 713.00	1185.11	9.2e-57	408	AM045329 UI-M-BH1-akn-h-11-0-0
939.est1:BI765023	+ 713.00	1187.39	2.2e-56	834	BI765023 603051106F1 NIH_MGC_11
940.est1:BI765023	+ 713.00	1177.98	2.3e-56	835	BI765023 601753388F1 NCI CGAP

901.est1:AV010326 - 706.50 1173.62 4.0e-56 461 1 AV010326 us95905.x1 NCI CGAP
902.est1:AA403935 + 706.00 1174.43 3.6e-56 397 1 AA403935 va32a03.r1 Guaywood
903.est1:BB218664 + 704.50 1166.61 9.8e-56 641 1 BB218664 BB218664 RIKEN full
904.est1:BI767700 + 703.50 1161.70 1.8e-55 858 1 BI767700 603060657F1 NIH_MGC
905.est1:BB747091 + 699.00 1160.52 2.1e-55 414 1 BB747091 BB747091 RIKEN full

seq_name: 901.est1:AV010326

seq_documentation block:

LOCUS AV010326 1039 bp mRNA linear EST 23-OCT-2001
DEFINITION AV010326 Mus musculus 18-day embryo C57BL/6J Mus musculus CDNA
clone 1110025H02, mRNA sequence.
ACCESSION AV010326
VERSION AV010326.2 GI:16356130
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M., and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

On May 11, 1999 this sequence version replaced gi:4787313.

Unpublished (2001)

CONTACT: Chile Owa

Genome Science Laboratory

RIKEN

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-9145

Fax: 81-298-36-9098

Email: genome-research@riken.go.jp

Carninci, P., Shibata, Y., Hayashizaki, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

Wadi, K., Fujiwara, S., Inoue, K., Togawa, Y., Itawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genomes 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamamaki, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., and Hayashizaki, Y.

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Location/Qualifiers

1.1039

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="1110025H02"

/clone_11b="Mus musculus 18-day embryo C57BL/6J"

FEATURES

Source

645 GCAGTGGCTGCAGCTCCAGACAGTGGAGCCCGGCCGATGACGCCCTCG 694

alignment_scores:

B1332055 697 bp mRNA EST 30-JUL-2001
 602982878F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5135755 5',
 mRNA sequence.
 B1332055
 B1332055.1 GI:15016712
 EST.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 697)
 NIH-NCBI <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM1133 row: k column: 20
 High quality sequence stop: 697.
 Location/Qualifiers
 1. 697
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5135755"
 /clone_lib="NCI CGAP_L19"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: liver; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.9 kb. Constructed by Life
 Technologies. Note: This is a NCI_CGAP Library."
 212 G 205 G 139 G

Ratio: 5.330 Gaps: 1
Percent Similarity: 99.138 Percent Identity: 98.707

Alignment block:
US-09-899-471-2 x B1332055

Align seg 1/1 to: B1332055 from: 1 to: 697

```

270 ArgAsnLeuTrpHisIleAlaArgLeuArgValLeuSerProGlyValTrp 286
|||||
2 AGAAGACCTCTGGACATAGCCAGGCTGGGCTGACTCTCCCAAGAGTATG 51
|||||
286 pGlnLeuAspAlaProCysCysLeuProGlyLysValThrLeuCysTrpG 303
|||||
52 GCACCTAGATGCGCTTGCTGCTGCTCCGGGCAAGTAACACTGTGCTGCG 101
|||||
303 InAlaProAspGlnSerProCysGlnProLeuValProProValProGln 319
|||||
102 AGGACCAAGACAGAGTCCCTGCCAGCAGCTGTGCTGACAGAGTCCAG 151
|||||
320 LysAsnAlaThrValAsnGlnProGlnAspPhe_GlnLeuValAlaGlyH 336
|||||
152 AAGAACCCACTGTGAATGAGCCACACAGATTCATGTTGGTGGCAGGCC 201
|||||
336 LsProAsnLeuCysValGlnValSerThrTrpGlnLysValGlnLeuGln 352
|||||
202 ACCCAACCTCTGTGTCCAGGTGAGCAGCTGGGAGAGTTCAGCTGCAG 251
|||||
353 AlaCysSerTrpAlaAspSerLeuGlyProPheLysAspAspMetLeuLe 369
|||||
252 GCGGTCTGTGTGGCTGACTCTCTGGGGCCCTTCAAGATGATGCTGTGT 301
|||||
369 uValGlnMetLysThrGlyLeuAsnAspThrSerValCysAlaLeuGlnP 386
|||||
302 AGTGGAGATGAAACCGGCTCAACACACATCAGTGTGCTGGTGGAGAC 351
|||||
386 roSerGlyCysThrProLeuProSerMetAlaSerThrArgAlaAlaArg 402
|||||
352 CCACTGGCTGTACACACACTGCCACACATGCTCCACGAGAGCTGCTGCC 401
|||||
403 LeuGlyGlnGluLeuLeuGlnAspPheArgSerHisGlnCysMetGlnLe 419
|||||
402 CTGGGAGAGGAGTGTCTGCAAGACTTCGATCACAACAGTATGAGAGCT 451
|||||
419 uTrpAsnAspAspAsnMetGlySerLeuTrpAlaCysPrometAspLysT 436
|||||
452 GTGGAGAGATACACATGAGGATGCTATGAGGCTGCCCATGAGACAGT 501
|||||
436 yrIleHisArgArgTrpValLeuValTrpLeuAlaCysLeuLeuAla 452
|||||
502 ACATCCACAGAGCGCTGGCTCTAGTATGGCTGGCTGCTACTCTTGCGT 551
|||||
453 AlaAlaLeuPhePhePheLeuLeuLysLysAspArgArgLysAlaAl 469
|||||
552 GCGGCGCTTTCTTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 601
|||||
469 aArgGlySerArgThrAlaLeuLeuLeuHisSerLysAlaAspGlyAlaGlyT 486
|||||
602 CCGTGGCTCCCGCAGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 651
|||||
486 yrGluArgLeuValGlyAlaLeuAlaSerAlaLeuSerGlnMet 500
|||||
652 ACAGAGCGTCTGTGGAGCAGCTGGGCTGCCGCTTGAAGCAGATG 695

```

seq_name: gb_est2:BG968034

seq_documentation_block:

LOCUS BG968034 784 bp mRNA linear EST 12-JUN-2001
DEFINITION 602832412F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4987035 5',
RNA sequence.
ACCESSION BG968034
VERSION BG968034.1 GI:14355671
KEYWORDS EST.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 784)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM10997 row: 9 column: 04
High quality sequence stop: 646.
Location/Qualifiers
1. 784
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4987035"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
173 a 231 c 229 g 151 t

FEATURES

source

BASE COUNT
ORIGIN

alignment_scores:

Quality: 1187.00 Length: 254
Ratio: 5.051 Gaps: 5
Percent Similarity: 92.520 Percent Identity: 89.370

alignment_block:

US-09-899-471-2 x BG968034

Align seg 1/1 to: BG968034 from: 1 to: 784

```

234 LeuAsnHisThrAspLeuValProCysLeuCysIleGlnValTrpSerLe 250
|||||
3 TTAACACACAGAGACTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 52
|||||
250 uGluProAspSerGluArgValGluPheCysProPheArgGluAspProg 267
|||||
53 AGAGCCAGACTCTGAGAGGCTGCAATTCCTCCCTCCGGAAGATCCCG 102
|||||
267 lYAlaHisArgAsnLeuTrpHisIleAlaArgLeuArgValLeuSerPro 283
|||||
103 GTGCACACAGAGACTCTGACACATAGCCAGGCTGGGCTGCTCTCCCA 152
|||||
284 GlyValTrpGlnLeuAspAlaProCysCysLeuPheGlyLysValThrLe 300
|||||
153 GGGGTATGGCAGCTAGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
|||||
300 uCysTrpGlnAlaProAspGlnSerProCysGlnProLeuValProProV 317
|||||
203 GTGCTGGCAGGACCAAGACAGAGTCTCTCCAGCCACTTGGCGCACAG 252
|||||
317 alProGlnLysAsnAlaThrValAsnGlnProGlnAspPheGlnLeuVal 333
|||||
253 TGCCCCAGAGAACGCCACTGTGAATGAGCCACACAGATTCAGATTGGTG 302
|||||
334 AlaGlyHisProAsnLeuCysValGlnValSerThrTrpLysValGln 350
|||||
303 GCAGGCGACCCCAACCTGTGTGTCCAGGTAGAGCACTGGGAGAGAGTTCA 352

```


350 nleuGlnAlaCysSerTrpAlaAspSerLeuGlyProPheLysAspAspM 367
 353 GCTCAAGACCGCTGCTTGGGCTGACTCCTTGGGGCCCTTCAAGAGATGA 402
 367 etleuLeuValGluMetLysThrGlyLeuAsnAspThrSerValCysAla 383
 403 TGCCTTAAAGTGGAGATGAACCCGCCCTCAACACACATCAGCTGTGCC 452
 384 leuGluProSerGlyCysThrProLeuProSerMetAlaSerThrArgAl 400
 453 TTGCAACCACTGCTGCTACACCTGCTCCACACTGAGACTCCACGAGAGC 502
 400 aalaArgLeuGlyGluGluLeuLeuGlnAspPheArgSerHisGlnCys 416
 503 TGCCTGCTGGAGAGAGTGTCTGCTCAGACTTCGATCCACACACTGA 552
 417 MetGlnLeuTrpAsnAspAspAspMetGlySerLeuTrpAlaCysProM 433
 553 ATGAGAGCTGTGGAAACGATGACACACTGGGATGCTATGGCCCTGCCCAT 602
 433 TAspLysTrpIleHisArgArgTrpValLeuValTrpLeuAla.CysLeu 449
 603 GGACAACTACTCCACAGCGCCCTGGGCTACTAGTATGGGCTGACTGA 652
 450 LeuLeu.Ala.....AlaAlaLeuPhePheLeuLeuLeuLysAla 464
 653 CTCCTGGGCTGACGGCGCATTAATCACTTCCCTCCTCAAGAAAGG 702
 464 spArgArg.LysAlaAlaArgGlySerArgTrpAlaLeuLeuHis 480
 703 ACCGACAGGAAAGCGCGCTGGGTCCGGAACGCGCAATTCCTCCCTC 752
 480 Iala 481
 753 CAGC 756
 eq_name: gb_est1:BB609618
 eq_documentation_block: 1006 bp mRNA linear EST 26-OCT-2001
 OCUS BB609618 RIKEN full-length enriched, 18 days embryo Mus musculus
 ERNITION CDNA clone 1110025H02 5', mRNA sequence.
 CCESSION BB609618
 ERSION BB609618
 EYWORDS BB609618.1 GI:16451334
 EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1006)
 AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
 Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda,
 M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,
 D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
 Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
 Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 Unpublished (2001)
 COMMENT
 TITLE Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9212
 Fax: 81-45-503-9216
 Email: genome-res@gs.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,
 M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)

wasgi,K., Fujisake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Natabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
 S., Kawai,U., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
 Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multichipillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
 K. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
 K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
 Hayashizaki,Y.
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 e mouse tissues.
 Location/Qualifiers
 1. 1006
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="1110025H02"
 /clone_lib="RIKEN full-length enriched, 18 days embryo"
 /dev_stage="18 days embryo"
 /lab_host="SOLR"
 /note="Site 1: XhoI; Site 2: SstI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5',
 GAGAGAGAGAGCGGCCCGCCACCTCCAGTTTCTTTTCTTTTNN 3'], cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. Second strand cDNA was prepared with the
 primer adapter of sequence [5',
 GAGAGAGAGAGATCCCAAGCTCATTAATTAATTAACCCCCCCCC 3'].
 cDNA was cleaved with XhoI and SstI."
 BASE COUNT 184 a 278 c 320 g 218 t 6 others
 ORIGIN
 alignment_scores:
 Quality: 1145.50 Length: 352
 Ratio: 4.227 Gaps: 12
 Percent Similarity: 76.989 Percent Identity: 69.886
 alignment_block:
 US-09-899-471-2 x BB609618 ..
 Align seg 1/1 to: BB609618 from: 1 to: 1006
 123 leuSerPheGlnAlaTyProIleAlaArgCysAlaLeuLeuGluValAlG 139
 14 CTCCTCTTCAGNCTACCCATCCCGCGCTGCTCTCTCTGAGAGTCCA 63
 139 nvalProAlaAspLeuValGlnProGlyGlnSerValGlySerAlaValP 156
 64 GGTCCCGCTGACCTGCTGCA.CCTGGTCAGTCCGTGGCTTCGCGGTAT 112
 156 heaSpCysPheGlnAlaSerLeuGlyAlaGluValGlnIleTrpSerTy 172
 113 TTGACTGTTTGGAGCTAGTCTTGGGCTGAGTACAGATGCTGGCTTCA 162
 173 ThrLysProArgTyGlnLysGluLeuAsnLeuTrpGlnIleLeuPro.. 188
 163 ACGAAGCCAGTACAGAAAGAGCTCAACCTCACACAGCACTGCTTA 212
 188


```

213 CTGACGGGGCTTGAAGTCCGGACAGCATCCAGACTGCTGGCTTCG 262
189 .....AspGlyAspAsnValLeuLeuThrLeuAsp 198
263 CTTGGCTCAATGTCTCTACAGATGGTGACAAATGCTCTTGACACAGCGGT 312
199 ValSerGluGluGlnAspPhe.SerPheLeuLeuTyrLeuArgProValP 215
313 GTCTCTGAGGAGCAGGCGCTTACGCTTCTTACTGTACCTGCGCCAGTCC 362
215 roAspAlaLeuLysSerLeuTyrTyrLysAsnLeuThrGlyProGlnAsn 231
363 CGGATGCTCTCAAACTTGTGTACAAAACCTGACTGAGCTCAGAAC 412
232 IleThrLeuAsnHisThrAspLeuValProGlyLeuGlyIleGlnValTr 248
413 ATTACTTTAAACACACAGACCTGGTCCCTGCTGCAATTCAGGCTG 462
248 pSerLeuGluProAspSerGluArgValGluPheCysProPheArgGluA 265
463 GTGCTGAGGAGCAGACTCTGAGAGGGTGCAATTCCTCCCGGAG 512
265 sPpGluGlyAlaHisArgAsnLeuThrPHisIleAlaArgLeuArgValLeu 281
513 ATCCGGGTGCACACAGGAGCCTCTGACATAGCCAGCGGTGCTACTG 562
282 SerProGlyValITrPGLnLeuAspAlaProCysCysLeuProGlyLysVa 298
563 TCCCGAGGGGATGAGCAGCTAGATCGCTTCTGCTCCCGGAGAGT 612
298 lThrLeuCysTrpGlnAlaProAspGlnSerProCysGlnProLeuValP 315
613 AACACTGTGCTGGCAGGCGCAGACAGATGCTCCCGCAGCTTGTG 662
315 roProValProGlnLysAsnAlaThrValAsnGluProGlnAspPheGln 331
663 CACACAGTCCCGCAGAGAGCCACCTGTGAATGAGCCACAGNATTTTCA 712
332 LeuValAlaGluHisProAsnLeuCysValGlnValSerThr.TrpGluL 348
713 TTGGTGGCAGGAGCCAGCCCTGTGTCTCCAGGTGAGACCTGGGGGA 762
348 yValGlnLeuGlnAlaCysSerTrpAlaAspSerLeuGlyProPheLys 364
763 GGGTTCAGTGCAGAGNGCTTGTG...GGCGACTCTTGGGGCTT..... 803
365 AspAspMetLeuLeuValGluMetLysThrGlyLeuAsnAsnThrSerya 381
804 TCAGGGGTGTATGCTGTAGTGAAGCGAGCGGCTCAGCACAGATCACT 853
381 lCysAlaLeuGluProSeryLysCysThrProLeuProSerMetAlaSer 397
854 CTGTGCTTGGGGCCAGCTGCTNTGACACATGCCNCAATNGCTCCACGGG 903
398 ThrArgAlaAlaArgLeuGlyGluLeuLeuGln.AsppheArgSeryH 414
904 ACT.....GCTGCTGAGAGAG.....TTGCTGAGAGCTTGCAATAAC 941
414 lsgInCysMetGlnLeuThrAsnAspAspAspMetGlySerLeuTrpAla 430
942 AT.....GGTTCACTGTGGACA 958
431 CysProMetAspLysTyrIleHisArgArgTrpValLeuValTrpLeu 446
959 TGC.....AGATGGATC...GTATGGCTC 979

```

```

ACCESSION      AI007139
VERSION        AI007139.1  GI:3216696
KEYWORDS
SOURCE
ORGANISM       Mus musculus
REFERENCE
AUTHORS        Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
                Geisel,S., Kucada,T., Lacy,M., Le,M., Martin,J., Morris,M.,
                Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
                Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilston,R. and
                Waterston,R.
TITLE          The WashU-HMI Mouse EST Project
JOURNAL        Unpublished (1996)
COMMENT        Contact: Marra M/Mouse EST Project
                WashU-HMI Mouse EST Project
                Washington University School of Medicine
                444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
                Tel: 314 286 1800
                Fax: 314 286 1810
                Email: mouseest@watson.wustl.edu
                This clone is available royalty-free through LIND; contact the
                IMAGE Consortium (info@image.llnl.gov) for further information.
                MGI:897119
FEATURES
source         High quality sequence stop: 495.
                Location/Qualifiers
                1..775
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone_image:1363899"
                /clone_lib="Soares_mammary_gland_NBMNG"
                /sex="male"
                /tissue_type="mammary gland"
                /dev_stage="4 weeks"
                /lab_host="DH10B"
                /note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia
                ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                primer [5',
                TGTTACCAATCTGAGTGGAGCGCGCCGAGATGTTTCTTTTCTTTTCTTTT
                T 3']; double-stranded cDNA was ligated to Eco RI
                adaptors (Pharmacia), digested with Not I and cloned into
                the Not I and Eco RI sites of the modified pT73 vector.
                RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                constructed and normalized by Bento Soares and M.Fatima
                Bonaldo."
BASE COUNT
ORIGIN         132 a      249 c      231 g      163 t
alignment_scores:
Quality: 1095.00      Length: 258
Ratio: 4.660          Gaps: 0
Percent Similarity: 91.085      Percent Identity: 86.822
alignment_block:
US-09-899-471-2 x AI007139
Align seg 1/1 to: AI007139 from: 1 to: 775
373 LysThrGlyLeuAsnAsnThrSeryValCysAlaLeuGluProSeryGlyC 389
|||||
3 AAAACCGCGCTCAACACACATCAGTCTGCTGTGAGAACCCAGTGGCTG 52
389 sThrProLeuProSerMetAlaSerThrArgAlaAlaArgLeuGlyGlu 406
|||||
53 TACACCTGCTGCGCAGCATGCGCTCCACGAGAGCTGCTCCGCTGGAGAG 102
406 lLeuLeuGlnAspPheArgSeryHisGlnCysMetGlnLeuThrPAsp 422
|||||

```


103 AGTGTGCAAGACTTCGATCACCAGCTGTATGACGCTGTGAACGAT 152
 423 ASPASmetGlySerLeuTrpAlaCysProMetAspLysTrpIleHisar 439
 153 GACACATGATGATCGCTATGGCGCTGCCATG. GACAAGTACATCCACAG 201
 439 GARGTTPValLeuValTrpLeuAlaCysLeuLeuLeuAlaAlaLeuP 456
 202 GCGCTGGGCTCTAGTATGGCTGGCTGCTACTCTTGGCTGGCGGCTTT 251
 456 hePhePheLeuLeuLeuLysLysAspArgArgLysAlaAlaArgLys 472
 252 TCTTCTCTCTCTCTTAAAGAGCCGAGAAAGCGCCGCTGCTCC 301
 473 ArgPheAlaLeuLeuLeuHisSerAlaAspGlyAlaGlyTrpIleArg 489
 302 CGCAGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 351
 489 uValAlaLeuAlaSerAlaLeuSerGlnMetProLeuArgValAla 506
 352 GGTGGGAGACATGGCGCTCCGCTTGAGCCAGATGCCCTGCGCTGCCG 401
 506 AlAspLeuTrpSerArgArgGluLeuSerAlaHisGlyAlaLeuAla 522
 402 TGGACCTGTGAGCCGCCGAGCTGAGCCGACGAGCCCTAGCCTGG 451
 523 PheHisGlnArgArgArgIleLeuGlnGlyValValIleLe 539
 452 TTACACCCACGAGGAGCCGCTATCTGACAGAGGCTGGCTGTAAATCT 501
 539 uLeuPheSerProAlaAlaValAlaGlnGlnGlnTrpLeuGlnLeu 556
 502 TCTCTTCTGCGCCGCGCGCTGGGAGGTCTACAGACTGCTGACTCTCC 551
 556 ltrhValGluProGlyProHisAspAlaLeuAlaAlaTrpLeuSerCys 572
 552 AGACAGTGTAGCCCGCGCCGATGAGCCCTCGCCCTTGCTAGCTGC 601
 573 ValLeuProAspPheLeuGlnGlnArgAlaIleGlyArgTrpValGly 589
 602 GTGCTTACCGATTTCTGCAAGC. CGGGGAGCCGTTAAACGTCGGGGGT 650
 589 lTrpPheAspGlyLeuLeuHisProAspSerValProSerProPheArg 606
 651 CTACTTGCAGCGGCGCTGCGCAACCAACTGTGTCTCTCCCGCTTCCG 700
 606 AlAlaProLeuPheSerLeuProTrpGlnLeuProAlaPheLeuAspAla 622
 701 GTGGCCCGTCTTCTCTCTGCTGCGCACTTGCAGCTTCTCTGAATGA 750
 623 LeuGlnGlyGlyCysSerThrSer 630
 751 CTTAAGGGAAGGTGATCAATTTCG 774

Seq name: gb_esc12:BF237179

seq documentation block:

LOCUS BF237179 722 bp mRNA linear EST 14-NOV-2000
 DEFINITION 60208133F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:4163217 5',

ACCESSION BF237179

VERSION BF237179.1 GI:1151096

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 722)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1AM9447 row: a column: 10
 High quality sequence stop: 584.
 Location/Qualifiers
 1. 722
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_id="NCI_CGAP_L19"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: liver; Vector: pCMV-SPORE6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.9 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 128 a 238 c 255 g 101 t

alignment_scores:
 Quality: 1058.00 Length: 207
 Ratio: 5.161 Gaps: 0
 Percent Similarity: 99.034 Percent Identity: 97.585

alignment_block:

US-09-899-471-2 x BF237179/rev ..

Align seg 1/1 to reverse of: BF237179 from: 1 to: 722

468 AlAlaArgGlySerArgTrpAlaLeuLeuHisSerAlaAspGlyAl 484
 629 GCGGCCGTGGCTCCGACAGACTGCTCTCTCGCTCCGCGGAGGAGC 580
 484 aGlyTrpGluArgLeuValGlyAlaLeuAlaSerAlaLeuSerGlnMet 501
 579 GGGCTACGAGCGCTGTGGGACACTGGCGCTCCGCTTACAGCCAGATGC 530
 501 rLeuArgValAlaValAspLeuTrpSerArgArgGluLeuSerAlaHis 517
 529 CACTGGCGGTGGCGCTGAGCCTGTGAGCCGCCGAGCTGAGCGCCAC 480
 518 GAlAlaLeuAlaTrpPheHisGlnArgArgIleLeuGlnGlu 534
 479 GAGGCCCTGACCTGGTTCCACACACAGCCGCGGTATCTGACAGAGGG 430
 534 yGlyValValIleLeuLeuPheSerProAlaAlaValAlaGlnGlnGln 551
 429 TGGCGGTAAATCTCTCTCTCTGCGCGCGCGCTGCGCAGTGTACG 380
 551 lTrpLeuGlnLeuGlnTrpValGluProGlyProHisAspAlaLeuAla 567
 379 AGTGGCTGAGCTCCAGACAGTGTGAGCCCGCGCGCTGACGCCCTGCGC 330
 568 AlATrPLeuSerCysValLeuProAspPheLeuGlnGlyArgAlaTrpG 584
 329 GCGTGGCTGAGCTGCTGTACCGATTTCTCTCAAGGCGCGGAGCCGG 280
 584 yArgTrpValGlyValTrpPheAspGlyLeuLeuHisProAspSerValP 601
 279 CCGGTACGTGGGGGTCTACTTGCAGGGGTGCGCACCCAGACTCTGTGC 230
 601 rSerProPheArgValAlaProLeuPheSerLeuProTrpGlnLeuPro 617
 229 CTTCCCGCTTCCGCGCGCGCTCTCTCTCTGCGCTGACAGCTGCGC 180
 618 AlApeLeuAspAlaLeuGlnGlyGlyCysSerThrSerAlaGlyArgPr 634

268 GCCCTCCCGCTCCGCGCCGCGCTCTTCCTCCCTCCGCGCAGCTGC 219

617 roAlaPheLeuAspAlaLeuGlnGlyGlyCysSerThrSerAlaGlyArg 633
|||||

218 CGGCTTTCCTGATGACTGCAGGAGAGCTGCTCCACTCCGCGGAGCA 169
|||||

634 ProAlaAspArgValGlyArgValThrGlnAlaLeuArgSerAlaLeuAs 650
|||||

168 CCCGCGGAGACCGGTGGAGACGAGTACCAGCGCGCTGCGCGCTGGA 119
|||||

650 pSerCysThrSerSerSerGluAlaAlaProGlyCysGluGluTrpAsp 667
|||||

118 CAGCTGACTCTTACCTCGGAAGCCCGGCTGCTCGAGGATGGGACC 69
|||||

667 euGlyProCysThrThrLeuGlu 674
|||||

68 TGGGACCCCTGCACCTACACTAGCA 46

seq_name: gb_est2:BE290565

seq_documentation_block:
LOCUS BE290565 562 bp mRNA linear EST 26-OCT-2000
DEFINITION 601089427F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:3484038 5',
mRNA sequence.
ACCESSION BE290565
VERSION BE290565.1 GI:9171696
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 562)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM8516 row: f column: 07
High quality sequence stop: 562.
Location/Qualifiers
1. 562
/organism="Mus musculus"
/strain="C57/B6"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt:
Site: 2; Notti: Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

BASE COUNT 86 a 190 c 180 g 106 t
ORIGIN

Alignment_scores:
Quality: 999.00 Length: 187
Ratio: 5.342 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:
US-09-899-471-2 x BE290565

Align seg 1/1 to: BE290565 from: 1 to: 562

414 HisGlnCysMetGlnLeuTrpAsnAspAspAsnMetGlySerLeuTrpAl 430
|||||

2 CACCAAGTGTATCAGCTGTGGAACGATGACACATGATGATGATGATGGC 51
|||||

430 acYsPromeTAspLysTyrIleHisArgArgTrpValLeuValTrpLeuA 447
|||||

52 CTGCCCATAGGCAACACTACATCCAGGCGCTGGTCTCTAGTATGGCTGG 101
|||||

447 IaCysLeuLeuLeuAlaAlaAlaLeuPhePheLeuLeuLeuLys 463
|||||

102 CTGCTACTCTTGGCTGGCGGCTTTCTTCTTCCCTCTTAATAAAG 151
|||||

464 AspArgValGlyAlaAlaAlaArgGlySerArgThrAlaLeuLeuHis 480
|||||

152 GACCCGAGGAAAGCGCGCGCTGCTCCGACGCGCTTCTCTCCACTC 201
|||||

480 rAlaAspGlyAlaGlyTyrGlnArgPheValGlyAlaLeuAlaSerAla 497
|||||

202 CCCCAGAGAGCGGCTACGAGCGTGTGGAGGACTGGCGTCCGCT 251
|||||

497 euSerGlnMetProLeuArgValAlaValAspLeuTrpSerArgArgGlu 513
|||||

252 TGAGCCAGATGCCACTCGCGCGTGGCGCTGGACCTGTGGACCCGCGGAG 301
|||||

514 LeuSerAlaHisGlyAlaLeuAlaTrpPheHisGlnArgArgArg 530
|||||

302 CTGACGCGGACGAGACCCCTACCTGTTCCACACAGGAGCGCCGTAT 351
|||||

530 eleuGlnGluGlyValValIleLeuLeuPheSerProAlaAlaValA 547
|||||

352 CTGGAGGAGGGGTGGCTGTGTAATCTTCTCTCTGCGCGCGCGCTGG 401
|||||

547 IaGlnCysGlnGlnTrpLeuGlnLeuGlnThrValGluProGlyProHis 563
|||||

402 CGCACTGTTCAGCAGTGGCTGCACCTCCACAGAGTGGAGCCGCGGCAT 451
|||||

564 AspAlaLeuAlaAlaTrpLeuSerCysValLeuProAspPheLeuGln 580
|||||

452 GACGCGCTCGCGCGCTGCTGACCTCGTGCATCCGATTTCTGCAAG 501
|||||

580 YargAlaThrGlyArgTyrValGlyValIlePheAspGlyLeuLeuHis 597
|||||

502 CCGGCGAGCGCGCGCTACGTGGGGTCTACTTCGAGCGGCGCTGCAAC 551
|||||

597 roAspSerVal 600
|||||

552 CAGACTCTGTG 562

seq_name: gb_est2:BF160122

seq_documentation_block:
LOCUS BF160122 635 bp mRNA linear EST 30-OCT-2000
DEFINITION 601767822F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3984149 5',
mRNA sequence.
ACCESSION BF160122
VERSION BF160122.1 GI:11040229
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 635)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
 Plate: ILAM9185 row: 1 column: 06
 High quality sequence start: 2
 High quality sequence stop: 632.
 Location/Qualifiers

1. 635

/organism="Mus musculus"
 /strain="C57/B6"
 /db_xref="taxon:10090"
 /clone_image="3984149"
 /clone_lib="NCI CGAP_Lu29"
 /tissue_type="spontaneous tumor, metastatic to mammary."
 Stem cell origin:
 /lab_host="DH10B"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site: 1: SalI; Site: 2: NotI; Cloned unidirectionally. Primer: Oligo df. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 92 a 229 c 204 g 110 t

alignment_scores:

Quality: 975.00 Length: 207
 Ratio: 4.924 Gaps: 4
 Percent Similarity: 95.652 Percent Identity: 93.237

alignment_block:

US-09-899-471-2 x BF160122

Align seg 1/1 to: BF160122 from: 1 to: 635

```

470 ArgGlySerArgThrAlaLeuLeuHisSerAlaAspGlyAlaGlyTyr 486
2 CGTGGCGCCGCGACGCTGCTCCACATCCGCGACGAGCGGCGCTA 51
486 rguArLeuValGlyAlaLeuAlaSerAlaLeuSerGlnMetProLeuA 503
52 CGAGCGCTGTGTAGGACACTGGCGCTCGCTGAGCCAGATGCCACTGC 101
503 rGValAlaValAlaSerLeuTrpSerArgGluLeuSerAlaHisGlyAla 519
102 GCGGCGCGCGGACGCTGTGAGCGCGCGCGAGCTAGCGGACGAGGCC 151
520 LeuAlaTrpPheHisSHSLnArgArgArgIleLeuGlnGlyGlyVa 536
152 CTAGCGCTGTTCACACACAGCGCGCGCTGATCTGACAGAGATGGCGT 201
536 lValIleLeuLeuPheSerProAlaAlaValAlaGlnGlnGlnGlnTrpL 553
202 GGTATATCTTCTCTGTGCGCGCGCGCGGAGTGTACAGCAGTGGC 251
553 eugInLeuGlnThrValGluProGlyProHisAspAlaLeuAlaTrp 569
252 TGCAGCTCCAGACAGTGGAGCGCGCTCGCATGACGCCCTCGCCCTGG 301
570 LeuSerCysValLeuProAspPheLeuGlnGlyArgAlaTrpGlyArgTy 586
302 CTCAGCTGCTGTACCCGATTTCTGCAAGCGCGGAGCGAGCGCGCTA 351
586 rValGlyValTyrPheAspGlyLeuLeuHisProAspSerValProSerP 603
352 CGTGGGGGTACTTCAAGCGGCTGTGACCCAGACACTGTGCTGCTGCC 401
603 rPheArgValAlaProLeuPheSerLeuProThrGlnLeuProAlaPhe 619
402 CGTTCGCGCGCGCGCGCTCTCTCTGCTGCTGCGACCTCCGCGCTTC 451
620 LeuAspAlaLeuGlnGlyGlyCysSerThrSerAlaGlyArgProAlaA 636
452 CTGGATGCACTGCAGGAGGCTGTCTCCACGTCGCCGGGGGAGCAGCGG 500

```

```

636 SPARGValGluArgValThrGlnAlaLeuArgSerAlaLeuAspSerC 652
501 ..CGGTGGACAGAGTGGCGCGCTGC...AGTCCCGCTGACACTG 545
652 sThrSerSerGluAlaProGlyCysGlnGlnTrpAspLeuGlyP 669
546 TACTCTGACTGCTGGAGCCCGAGGCTGTGCGAGATGAGTGGACCTGGAC 595
669 rOCysThrThrLeuGlu 674
596 CTGCACTACACTACAGTA 612

```

seq_name: gb_est2:BI251117

seq_documentation_block:

LOCUS BI251117 839 bp mRNA linear EST 17-JUL-2001
 DEFINITION 602994315F1 NCI_CGAP_Mam5 Mus musculus CDNA clone IMAGE:5150378 5',
 mRNA sequence.
 ACCESSION BI251117 GI:14800198
 VERSION BI251117.1 GI:14800198
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 839)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: ILAM1371 row: m column: 03
 High quality sequence stop: 836.

FEATURES

source

1. 839

/organism="Mus musculus"
 /strain="C57/B6"
 /db_xref="taxon:10090"
 /clone_image="5150378"
 /clone_lib="NCI_CGAP_Mam5"
 /tissue_type="tumor, gross tissue"
 /dev_stage="7 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI; Site: 2: NotI; Cloned unidirectionally. Primer: Oligo df. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"

BASE COUNT 146 a 248 c 268 g 177 t

alignment_scores:

Quality: 967.50 Length: 211
 Ratio: 5.013 Gaps: 2
 Percent Similarity: 91.469 Percent Identity: 91.469

alignment_block:

US-09-899-471-2 x BI251117

Align seg 1/1 to: BI251117 from: 1 to: 839

```

1 MetProValSerTrpPheLeuLeuSerLeuAlaLeuGlyArgAspProVa 17
199 ATGCTGTGTCTGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 248

```


17 lvalvalserleuGlualgLeuMetGluProGlnAspThrAlaArgCys 34
 249 GGCTGCTCTCTGAGAGACTGATGGAGCCCTCAGACACTCAGCTCT 298
 34 erleuGlyLeuSerCysHisLeuTrpAspGlyAspValLeuCysLeuPro 50
 299 CTCTAGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 348
 51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuG1 67
 349 GGAAGCTCCAGCTGCTGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCA 398
 67 nThrGluLeuValLeuArgCysProGlnLysThrAspCysAlaLeuArgV 84
 399 GACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 448
 84 AlArgValValValHisLeuAlaValHisGlyHisThrAlaGluProGlu 100
 449 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAA 498
 101 GluAlaGlyLysSerAspSerGluLeuGlnGluSerArgAsnAlaSerle 117
 499 GAAAGCTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 548
 117 uGlnAlaGlnValValLeuSerPheGlnAlaArgProIleAlaArgCysA 134
 549 CCAGGCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598
 134 lAlaLeuGlnValAlaGlnValProAlaAspLeuValGlnProGlyGlnSer 150
 599 CCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 648
 151 ValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlnValGluVal 167
 649 GTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 698
 167 lGlnIleTrpSer.TyrThrLysProArgTyrGlnLysGluLeuAsnLeu 183
 699 ACAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 748
 184 ThrGlnGlnLeuProAsp..... 189
 749 ACACAGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798
 190GlyAspAsnValLeuLeu 195
 799 AGCTGCTGGGATGTGACATGCTCTCTG 829

Seq_name: gb_est2.BG761977

Seq documentation block:
 LOCUS BG761977 935 bp mRNA linear EST 15-MAY-2001
 DEFINITION 60271892881 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4858806 5',
 mRNA sequence.
 ACCESSION BG761977
 VERSION BG761977.1 GI:14072630
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 935)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC/DCFD/DTF
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LNCM713 row: h column: 07
 High quality sequence stop: 813.
 Location/Qualifiers
 1.935
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4858806"
 /clone_1lb="NIH_MGC_49"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 154 a 301 c 291 g 189 t

ORIGIN

alignment_scores:
 Quality: 912.50 Length: 307
 Ratio: 3.755 Gaps: 8
 Percent Similarity: 79.153 Percent Identity: 62.866

alignment_block:
 US-09-899-471-2 x BG761977

Align seg 1/1 to: BG761977 from: 1 to: 935

237 ThrAspLeuValProCysLeuGlnIleGlnValTrpSerLeuGluProAs 253
 2 ACAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 51
 253 pSerGluArgValGluPheCysProPheArg.GluAspProGlyAlaHis 269
 52 CTCGTTAGGACGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 101
 270 ArgAsnLeuTrpHisIleAlaArgLeuArgValLeuSerProGlyValTr 286
 102 CAGAACCTCTGCAAGCCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 151
 286 pGlnLeuAspAlaProCysCysLeuProGlyLysValThrLeuCysTrpG 303
 152 GCTGCTGAGGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 201
 303 lAlaLProAspGlnSerProCysGlnProLeuValProProValProGln 319
 202 GGGCTCCGGGGGGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 251
 320 LysAsnAlaThrValAsnGluProGlnAspPheGlnLeuValAlaGlyTr 336
 252 GAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 298
 336 sProAsnLeuCysValGlnValSerThrTrpLysValGlnLeuGlnA 353
 299 CCTTAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 348
 353 lAcysSerTrpAlaAspSerLeuGlyProPheLysAspAspMetLeuLeu 369
 349 AGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 398
 370 ValGluMetLysThrGlyLeuAsnAsnThrSerValCysAlaLeuGlnTr 386
 399 TTGGAGACAGAGGCCCCCAGACAGATCCCTGCTGCTGCTGCTGCTG 448
 386 oSerGlyCysThrProLeuProSerMetAlaSerThrArgAlaAlaArgL 403


```

449 CAGTGGCTTACTTACTACCAACAGCAAGCCTCCAGAGGCGCTCGCC 498
403 euGlygluLeuLeuGlnAspPheArgSerHisGlnCysMetGlnLeu 419
499 TTGGAGAGTACTTACTACAGACCTGCAGCGCAGTGTCTCAGCTA 548
420 TTPAsuAspAspAspMetGlySerLeuTPrLacysPrometAspLysTy 436
549 TGG...GACGATGACTTGGAGCGCTATGGCGCTCCCGCCATGACAAATA 595
436 rllhAsArgArgTPrValLeuValTPrLeuAlaCysLeuLeuAla 453
596 CATCCACAAGCCCTGGCGCCCTGTGGCTGGCGCTCTACTCTTGGCG 645
453 lAlaLeuPhePhePheLeuLeuLysLysAspArg... 466
646 CTGGCGCTTCCCTCATCTCTCTCTCAAAAAGATCACCGCAAGGCTT 695
467 .....LysAlaAlaArgGly.....Se 472
696 GCGTTGAGGCTCTGAACACAGACGCTCCGCTGGCGCGCCGCCAG 745
472 rArgTrAlaLeuLeuLeuHisSerAlaAspGlyAlaGlyTyrGluArgL 489
746 GCCGCGCGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 795
489 euVal.GlyAlaLeuAlaSerAla.LeuSerGlnMetProLeuArgVal 505
796 TGGTGGGGCGGCTGTGGCG...TCGCGTCCAGCTTGGCGCTGGCGTGG 842
505 lAvAlaPLeuTPrSerArgArgGluLeuSerAlaHisGlyAlaLeuAla 521
843 CCGTACACTGTTCAGCGCGTGTGACTGACGCGCA.GGGGCGCTGGCT 891
522 TTPheHisHisGln 526
892 TGGTTCCCGCGCAG 906

Seq.name: gb_est2:BI282622

Seq_documentation_block:
LOCUS BI282622 619 bp mRNA linear EST 19-JUL-2001
DEFINITION UI-R-CM05-cgd-b-11-0-UI.s1 UI-R-CM05 Rattus norvegicus cDNA clone
ACCESSION BI282622
VERSION BI282622.1 GI:14933575
KEYWORDS EST.
SOURCE Norway rat.
SDRCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 619)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704447
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@iuii.uwee.uia.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide Poly A
tail. The sequence tag present in the cDNA between the Poly A
and the oligo-dT track served to identify it as a clone from the
non-normalized rat aorta pool library cDNA library Preparation:
M.B. Soares Lab Clone distribution: clones will be available
through Research Genetics (www.resgen.com)

```

```

Seq primer: M13 forward
POLYA=Yes.
FEATURES
    source
        Location/Qualifiers
            1..619
                /organism="Rattus norvegicus"
                /strain="Sprague-Dawley"
                /db_xref="taxon:10116"
                /clone="UI-R-CM05-cgd-b-11-0-UI"
                /clone_11b="UI-R-CM05"
                /dev_stage="ADULT"
                /lab_host="DHI0B (Life Technologies)"
                /note="Vector: pT73D-Pac (Pharmacia) with a modified
                polylinker. Site_1: Not I. Site_2: Eco RI. The UI-R-CM05
                library is a non-normalized library constructed from the
                following rat tissues: embryonic day 19, embryonic
                day 21, adult day 1, adult day 12, adult day 75, adult day
                200. For a detailed description of the library from which
                this clone was derived, please visit our web site at
                ratest.eng.uiowa.edu. The subtraction has been previously
                described in (Bonaldo, Lennon and Soares, Genome Research
                6:791-806, 1996)
                TAG_L1B-UI-R-CM05
                TAG_TISSUE-rat aorta pool
                TAG_SEO-CTGTAGATC"
BASE COUNT 105 a 197 c 211 g 106 t
ORIGIN

alignment_scores:
    Quality: 893.00 Length: 193
    Ratio: 4.907 Gaps: 0
    Percent Similarity: 94.301 Percent Identity: 89.119

alignm_block:
US-09-899-471-2 x BI282622/rev ..
Align seg 1/1 to reverse of: BI282622 from: 1 to: 619

482 AspGlyAlaGlyTyrGluArgLeuValGlyAlaLeuAlaSerAlaLeu 498
619 GATGAGACATACATACAGACGCTGTGGCGCTCGGCTCGGCTTGGAG 570
498 rGlnMetProLeuArgValAlaValAspLeuTPrSerArgArgGluLeu 515
569 CCAGATGCACCTGGCGCTGGCGCTGGAGCTGTGGAGCGCGCGAGCTGA 520
515 eAlaHisGlyAlaLeuAlaTPrPheHisHisGlnArgArgGlyLeu 531
519 CCGCGCAGGAGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGAG 470
532 GlnGlnGlyValValAlaLeuLeuPheSerProAlaAlaValAlaG1 548
469 CAGGAGATGGCGGTGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
548 ncysGlnGlnTPrLeuGlnLeuGlnThValGluProGlyProHisAsp 565
419 GTGTGACAGTGGCTGCAGCTCCAGACGCTGAGCGCGGCGGATGAG 370
565 lAlaValAlaTPrLeuSerCysValLeuProAspPheGlnGlnArg 581
369 CTGTGGCGCGCTGTGATGCTGTCTGTCTGTCTGTCTGTCTGTCTGT 320
582 AlAThGlyArgTyrValGlyValTyrPheAspGlyLeuHisAsp 598
319 GCGAGCGCGCGCTACGTCGGGGCTTACTTTCAGCGGCTGTGACACCA 270
598 pSerValProSerProPheArgValAlaProLeuPheSerLeuProThr 615
269 CACTGTACCGCGCTGTCTCGCGCGCGCGCTCTCTCTCTCTCTCTCT 220
615 lIneuProAlaPheLeuAspAlaLeuGlnGlyGlyCysSerThSerAla 631
219 AGCTGGCGCTTCTCTGATGCTGACGAGGAGACAGCTCCAGTCCGCC 170

```


632 GLYRPRGALAAAPARVAlGILURPVSLTHGALNALeaurgsral 648
169 GGGGGACCTGGCGAACGGAGGAGAACGAGTGGCCAGGCCCTGGTTCGC 1200
648 ALeuAsPserCysThrSerSerSergIuLapProGlyCysCysGluGur 665
119 CCTRAGACAGCTGCTACTTAGAGCTCGGAAACCCCGAGCGCTGCCGAGGAAT 70
665 rPAsPLeuGlyProCysThrThrLeuGlu 674
69 GGGAGCTGGAGACCTCGCCCTGCAGCTCGAA 41

seq_documentation_block;

LOCUS	561 bp	EST 31-JAN-2002
DEFINITION	musculus cdna clone C0660H10 3'	musculus cdna clone C0660H10 3'
ACCESSION	BM2102318	BM2102318

BASE COUNT ORIGIN	101 a	177 c	201 g	82 t
----------------------	-------	-------	-------	------

alignment_scores:

Quality:	876.00	Length:	180
Ratio:	4.977	Gaps:	0
Percent Similarity:	97.778	Percent Identity:	96.667

alignment_block:
US-09-899-471-2 x BM210218/rev

Align seg 1/1 to reverse of: BM210218 from: 1 to: 561

```

495 SerAlaIleuSerGlnMetProLeuArgValAlaValaspLeutripSerAr 511
|||||::: |||||
500 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG

```

511 gargluleuserAlahisglYalaleualatrpphenishisglNargA 528

528 rgargIleleuGlnGlyValValIleleuLeuPheSerProAla 544

[illegible]

411 GCCGTGCGCAGTGTACGAGTGGCTGCAGCTCCAGACAGTGGAGCCCGG 362

361 GCGGATGACGCCCTGGCTAGCTGCCTACCCGATTTC 317

5/8 euglnclyarGalatnrglyarqlyrvalgllyvallyrphneaspqlyleu 594

595 leuH1sPrAspServa]proSerProphEArvA]A]aprolEuphSe 611

[illegible]

211 CCTGCCTCGCAGCTGCCGCTTCTCGATGCACTGCAGGAGGCTGCT 162

020 6111 Bernadette Lynn G. Christiansburg v. J. Olin Williams 04

seq_name: qb_est2:BM009552

```
seq_documentation_block
```

ACCESSION	U000052
DEFINITION	603630118F1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5443900 5', mRNA sequence.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 702)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LNCM1922 row: C column: 05
High quality sequence stop: 697.
Location/Qualifiers

FEATURES

source
1. 702
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1b="NIH-MGC-41"
/issue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC library."

BASE COUNT 118 a 218 c 222 g 144 t
ORIGIN

alignment_scores:

Quality:	872.50	Length:	232
Ratio:	4.256	Gaps:	1
Percent Similarity:	88.362	Percent Identity:	71.983

alignment_block:
us-09-899-471-2 x BM009552 ..

Align seg 1/1 to: BM009552 from: 1 to: 702

```
323 ThValasngluproglaspheglneuvAlaIaIyHISProasne 339
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 ACTGTGACAAAGTTCTCGAGTTCACATGCTGCAAGGCCACCTAAGCT 55
339 uCysValaIglValSerThrtIpGluLysValaIglneuglAlaIaCysSert 356
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
56 CTGTGTTCAAGTGAACAGCTCGAAGAGCTGCAGAGCTGCAGAGCTTCT 105
356 rPalAaSPserleuGlyProphelysAspAspMetleuLeuValaIuMet 372
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
106 GGCCTGACTCCTCGGCTCTCAAGACGATGCTACTGTTGGAGACA 155
373 LysThrcIyLeuAsnAsnThSerValaIaIeugluproSergIyCY 389
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
156 CGAGGCCCCAGGACAAACAGATCCCTGTCCTTGGAACCCAGTGGCTG 205
389 sThrProleuProserMetAlaSerThrAlaIaIaArgleuglyIug 406
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
206 TACTTCACTACCAAGCCCTCCAGAGCGCAGCTCGCCTTGAGAGAT 255
406 lueLeuGlnAspPheArgSerHisGlnCysMetGlnleuTripsnAsp 422
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
256 ACTTACTACAAAGACTCGAGTCCAGGCCAGTGTCTGCAGCTATGG...GAC 302
423 AspAsnMetClySerleuTripsnAlaCysProMetAspLysTYrIleHisAr 439
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
303 GATGACTTGGAGAGCGCTATGGCGCTGCCCATATGACAAATACATCCACAA 352
439 GARGTrpAlleuValTrpLeuAlaCysleuLeuAlaAlaIaIeup 456
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
353 GCGCTGGGCGCTCGGTGTGGCTGGCCCTACTCTTGGCCGCTGGCTTT 402
456 hePhePheleuLeuLysLysAspArgArgLysAlaIaIaArgLysSer 472
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
403 CCTTATCTCTCTTCANAAAGATACGCCGAAAGCGGCCGAGGGGC 452
473 ArgTrpAlaIaIeuleuLeuHisSerAlaAspGlyAlaGlyTrpLuarGle 489
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
453 GCGCGGCTCTGCTCTCTACACAGCCAGTACTCGGGTTTCAGAGCCCT 502
489 uValaIyAlaIeuaIaSerAlaIeuserglmetProleuArgValaIaIay 506
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
503 GGTGGGCGCCCTGGCGGTGGCCCTGTGACAGTGCCTGCGGCTGGCCG 552
506 AlAspLeuTripsnArgArgGluLeuSerAlaHisGlyAlaIaIaIaTrp 522
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
553 TAGACCTGTGAGCCGTCTGTAAGTGAAGCGGAGGCGCGGCTGGCTTGG 602
523 PheHisGlnArgArgArgLyleuGlnGlyValaIaIaIeIe 539
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
603 TTTCACGCGGACGCGGCCAGAACTGCAAGAGGCGGCGTGGTCTT 652
539 uLeuPheSerProAlaIaIaValaIaGlnCysGlnGlnTrpLeuGln 554
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
653 GCTCTTCTCTCCGGTGGCGGTGGCGGTGACAGGAGTGGCTACAG 698
```


Job of: US-09-899-471-2_COPY_21_440 to: GenEmbl:* out_format : pfs

Date: Sep 29, 2002 1:03 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=ifamer_p2n.model -DEV=x1h
-O=/cgn2.1/USPOT/US09899471/runat_27092002_142439_17908/app-query.fasta.1.2519
-DB=GenEmbl -OEMT=fastmap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEC=0.000 -LOOPEXT=0.000 -OGAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FEAPOP=6.000
-FEAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELAP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blotsun62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OEMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=2000000000 -USER=US09899471.ecgn1.1.11274
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPMX -WAIT -THREADS=1

Search information block:

Query: US-09-899-471-2_COPY_21_440
Query length: 420
Database sequences: 1797656
Database length: 187333701
Search time (sec): 11857.370000

sequence	Strd Orig	zscore	EScore len	Documentation
b.ro:BC004759	+	2240.00	3193.34	1.7e-169 2015
b.pat:AX350976	+	1720.00	2461.29	1.9e-129 2094
b.pat:AX350978	+	1720.00	2461.29	1.9e-129 2094
b.pat:AX073993	+	1554.00	2210.05	9.7e-115 2255
b.pat:AX092430	+	1554.00	2209.56	1.0e-114 2380
b.pat:AX180776	+	1554.00	2209.56	1.0e-114 2380
b.pat:BC006411	+	1554.00	2209.10	1.1e-114 2507
b.pat:AX350973	+	1465.00	2082.40	1.3e-107 2308
b.pat:AX136341	+	1397.00	1994.32	3.6e-102 2499
b.pat:AX073996	+	1154.50	1640.26	5.3e-83 2076
b.pat:AX073995	+	1148.00	1629.43	2.1e-82 2076
b.pat:AX350975	+	1076.00	1526.20	1.2e-76 1725
b.pat:AX073998	+	852.50	1207.96	6.4e-59 1725
b.pat:AX136589	+	464.00	659.82	2.2e-28 694
b.pat:AC026196	+	316.00	337.57	8.8e-14 191834
b.pat:AC018809	+	316.00	336.47	1.0e-13 216905
b.pat:AC024612	+	310.00	390.24	2.3e-13 166620
b.pat:AC007783	+	310.00	390.13	2.3e-13 168665
b.pat:AC022007	+	310.00	389.86	2.4e-13 173765
b.pat:AC026685	+	309.50	380.74	2.4e-13 145435
b.pat:AC024612	+	307.50	366.66	3.6e-13 166620
b.pat:AC068099	+	320.00	261.85	3.2e-06 157773
b.pat:AC096599	+	204.00	246.24	2.4e-05 69815
b.pat:AC068099	+	176.00	198.85	0.1036 157773
b.pat:AC026685	+	174.00	196.72	0.0136 145435
b.pat:AC026196	+	154.00	165.00	0.7339 191834
b.pat:AC030982	+	151.00	202.11	0.0068 2012
b.pat:AF356232	+	144.50	185.17	0.0596 4715
b.pat:AF180778	+	137.50	182.23	0.0869 2138
b.pat:HSB845547	+	128.50	132.09	53.96 137139
b.pat:AC099658	+	126.50	124.15	0.49.47 241861
b.pat:AF130972	+	123.00	169.46	0.4474 1206
b.pat:AC093640	+	124.00	136.12	32.17 42559
b.pat:AC093640	+	123.50	120.99	224.16 213045
b.pat:AC092466	+	123.50	120.85	228.22 216408
b.pat:AC010973	+	123.50	120.59	235.74 222605
b.pat:AF19049	+	122.50	163.92	0.9096 1500
b.pat:RRE65	+	122.50	160.30	1.45 2249
b.pat:AF333983	+	122.50	159.16	1.68 2554
b.pat:AC007603	+	122.00	119.54	269.74 196924

gb.ro:AF206720 + 120.00 155.42 2.71 2601 AF206720 Mus musculus Fe65
gb.htg:AC103308 + 120.00 117.57 347.58 178349 AC103308 Rattus norvegicus
gb.htg:AC060230 - 119.00 110.57 852.22 331909 AC060230 Homo sapiens chr
gb.htg:AC073808 + 118.50 116.98 374.51 149727 AC073808 Mus musculus c1c
gb.pr:AC073957 - 118.50 114.56 510.83 196204 AC073957 Homo sapiens BAC

seq_name: gb.ro:BC004759

seq_documentation_block:

LOCUS BC004759 2015 bp mRNA linear ROD 12-JUL-2001
DEFINITION Mus musculus, clone MGC:6973 IMAGE:3154616, mRNA, complete cds.
ACCESSION BC004759
VERSION BC004759.1 GI:13435809
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2015)

REFERENCE 1 (bases 1 to 2015)

AUTHORS Villalon, D.R., Luna, R.A., Hale, S.M., Huylk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

REMARK COMMENT
Submitted (21-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalon@bcm.tmc.edu
Villalon, D.R., Luna, R.A., Hale, S.M., Huylk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.liml.gov>
Series: IRAC Plate: 6 Row: C Column: 3.

FEATURES

1. 2015
location/qualifiers

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="MGC:6973 IMAGE:3154616"
/tissue_type="Mammary tumor. Breast./EL: MMTV-Cre model. 10 months old, gross tissue."
/clone_lib="NCI CGAP Mam3"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
177. 1880
/codon_start=1
/product="Unknown (protein for MGC:6973)"
/protein_id="AAH04759.1"
/db_xref="GI:13435810"
/translation="MPVSMFLSLSLALGSRNPVVSLERLMEPODTARCSGLSCLHMDG
DVLCPSLSIAGPVPVLPTRLOPELVBCPQKTCACLRVYVHIAVGHNAPEPEA
GKSPATELQNSRNLSAQVYVSPQAYPARCLLEVVQPADLVQSGSAGVDFCE
ASLGAETQINSYTPRQKEKLNTDQLDPCROLEVRDSIOSCVLPLWLVSDGDNVL
LTLDSSEQDSFLLYLRPVPDAKSLWYKNTIGQNTTNTDLPVCLCIOWSLEP
DSERVECPREFEDGAGHNMHIAIRLVLSYGVMDLAPCCIGKVTCLGWAPODSSPC
OPLVPPVQKNAATVNEPODFOLVGHVPHICVQVSWEKVQVQACIADSLGAPDQML
LVEKKTGLNNTSVCLALPESGCTPLPSMASTAAIRGEFLVDPRCHOCQMLNDNDMG
RPAVSGSTSTGCTOTLCPRPSASPRNSPCPRCSRLSMHMCREAPLPRGDPRTGWN
E"

CDS

BASE COUNT 379 a 616 c 595 g 425 t
ORIGIN

alignment_scores:

Quality: 2240.00 Length: 444
 Ratio: 5.359 Gaps: 1
 Percent Similarity: 94.144 Percent Identity: 94.144

alignment_block:

US-09-899-471-2_COPY_21_440 x BC004759

Align seg 1/1 to: BC004759 from: 1 to: 2015

```

1 LeuGluArgLeuMetGluProGlnAspThrAlaArgCysSerLeuGlyLe 17
237 CTGGAGAGACTGATGAGAGCTCAGAGACTGCACGCTGCTTAGGGCT 286
17 USerCysHisLeuTrpAspGlyAspValLeuCysLeuProGlySerLeu 34
287 CTCTGCGACCTGTGGATGTGACGTCTGCTGCTGAGAGCTCC 336
34 InSerAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeu 50
337 AGTGTGCCCCAGGCCCTGTGCTAGTGCCTACCCGCTGCAGAGAGCTG 386
51 ValLeuArgCysProGlnLysThrAspCysAlaLeuArgValArgVala 67
387 GTCTGAGGTGTCCACAGAGACAGATTGCGCCCTGTGCTGCTGCTGCT 436
67 ValHisLeuAlaValHisGlyHisThrAlaGluProGluGluAlaGlyL 84
437 GGTCCACTGGCCCGTCAGTGGGCGACTGGGCGAGAGCTGAAGAGCTGGAA 486
84 YSerAspSerGluLeuGlnGluSerArgAsnAlaSerLeuGlnAlaGln 100
487 AGCTGTGATTCAGAACTCCAGAGAGTGTAGAGAGCCCTCTCCAGGCCAG 536
101 ValValLeuSerPheGlnAlaThrProIleAlaArgCysAlaLeuLeu 117
537 GTGTGCTCTCTCTCCAGGCCATCCAGCCGCTGTGCTGCTGCTGCTGGA 586
117 ValGlnValProAlaAspLeuValGlnProGlyLysSerValGlySera 134
587 GGTCCAGAGTCCCGCTGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 636
134 ValAlaPheAspCysPheGlnAlaSerLeuGlyAlaGluValGlnIleTrp 150
637 CGGATTTGAGCTGTTTGAAGGCTACTCTTGGGGCGTACAGATCTG 686
151 SerTrpThrLysProAlaGlyGlnLysLeuLeuAsnLeuThrGlnGlnLe 167
687 TCCATACAGAGAGCCAGCTACAGAGAGAGCTCAACTCAGACAGAGAGCT 736
167 UPTo..... 168
737 GCTGACTGCAAGGGGCTTGAAGTCCGGGAGAGCATCCAGAGCTGCTGGG 786
169 .....AspGlyAspAsnValLeuLeuThr 176
787 TCTGCGCCTGGCTCAATGTGTCTACAGATGTGACATCTCTTCTGACA 836
177 LeuAspValSerGluGluGlnAspPheSerPheLeuLeuTrpArgPr 193
837 CTGATATCTCTGAGGAGCAGAGACTTACTTACTGACTGCTGCTGCTGCT 886
193 OValProAspAlaLeuLysSerLeuTrpTrpLysAsnLeuThrGlyPro 210
887 AGTCCCGAGAGCTCTCAAAATCTGTGTACAAAAACCTGACTGAGCTTC 936
210 InAsnIleThrLeuAsnHisThrAspLeuValProCysLeuCysIleGln 226
937 AGAACAATTACTTAAACACACAGAGCTGTGCTGCTGCTGCTGCTGCTG 986
227 ValTrpSerLeuGluProAspSerGluArgValGluPheCysProPheAr 243
987 GTGTGTGCTAGAGCCAGACTGTGAGAGGGTGCAGAAATCTGCCCCCTTCG 1036

```

```

243 gclUsApProGlyAlaHisArgAsnLeuTrpHisIleAlaArgLeuArgv 260
1037 GGAGATCCCGGTGACACAGAGAACTGTGACATAGCCAGCGCTGGCG 1086
260 allUserProGlyValTrpGlnLeuAspAlaProCysCysLeuProGly 276
1087 TACTGTCCCGAGGGGTATGGACACTAGATGCGCTTGTCTGTGCGGGC 1136
277 LysValThrLeuCysTrpGlnAlaProAspGlnSerProCysGlnProle 293
1137 AAGTATACCTGTGCTGGCAGGACAGACAGATGCTGCTGCGGACT 1186
293 uValProProValProGlnLysAsnAlaThrValAsnGluProGlnAsp 310
1187 TGTGCCACCGAGTCCCGAGAGAACCCACTGTGAATGAGCCACAGATT 1236
310 heGlnLeuValAlaGlyHisProAsnLeuCysValGlnValSerThrTrp 326
1237 TCCAGTGTGGGCGAGGCCACCCCAACCTCTGTGTCCAGGTGAGCAGCTG 1286
327 GluLysValGlnLeuGlnAlaCysSerTrpAlaAspSerLeuGlyProp 343
1287 GAGAGAGTTCACCTGCAACGCTCTGTGCTGCTGCTGCTGCTGCTGCT 1336
343 elYsAspAspMetLeuLeuValGluMetLysThrGlyLeuAsnAsnThr 360
1337 CAAGATGATATGCTGTGTAGTGAGATGAAACCGGCTCAACACAT 1386
360 eValCysAlaLeuGluProSerGlyCysThrProLeuProSerMetAla 376
1387 CAGTGTGCTGCTGGAACCCAGGTGCTGTACACACTGCTGCCACATGCGC 1436
377 SerThrArgAlaAlaArgLeuGlyGluGluLeuLeuGlnAspPheArgSe 393
1437 TCCAGAGAGCTGCTGCTGCTGCGAGAGAGTGTCTGACAGATCTCCATC 1486
393 rHisGlnCysMetGlnLeuTrpAsnAspAspAsnMetGlySerLeuTrpA 410
1487 ACACCAAGTATGACAGCTGTGAGACGATGACACATGAGATGCTATGGG 1536
410 lacYsProMetAspLysTrpIleHisArgArg 420
1537 CTTGCTCCATGGACATGATCATCCAGAGCGC 1568

```

seq_name: gb_pat:AX350976
 seq_documentation_block:
 LOCUS AX350976 2314 bp DNA linear PAT 06-FEB-2002
 DEFINITION Sequence 10 from Patent WO0190358.
 ACCESSION AX350976
 VERSION AX350976.1 GI:18616352
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (sites)
 AUTHORS Gorman,D.M.
 TITLE Mammalian receptor proteins; related reagents and methods
 JOURNAL Patent: WO 0190358-A 10 29-NOV-2001;
 SCHERING CORPORATION (US)
 FEATURES
 source location/Qualifiers
 1..2314
 /organism="unidentified"
 /db_xref="taxon:32644"
 /note="rodent; surmised Mus musculus"
 199..2295
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD22836.1"
 /db_xref="GI:18616353"
 /translation="MYSWRLLSLALGRNPVVSLERLMERQDTRARSLGSLHMDG
 DVLCPLGSLQASAPGPIVIVPRLQTEIVLRCKQKTDCAICVRYVHLAVHGMAREPEEA

mat_peptide
 /product="unnamed"
 259. .2292
 411 a 725 c 705 g 473 t

BASE COUNT 411 a 725 c 705 g 473 t
 ORIGIN

alignment_scores:
 Quality: 2240.00 Length: 444
 Ratio: 5.359 Gaps: 1
 Percent Similarity: 94.144 Percent Identity: 94.144

alignment_block:
 US-09-899-471-2_COPY_21_440 x AX350976 ..

Align seg 1/1 to: AX350976 from: 1 to: 2314

```

1 LeuGIuArgLeuMetGluProGluAspThrAlaArgCysSerLeuGlyLe 17
259 CTGGAGAGATGATGAGACCTCAGACACTGACGCTCTCTAGCGCT 308
17 uSerCysHisLeuTrpAspGlyAspValLeuCysLeuProGlySerLeu 34
309 CTCCTGCACCTCTGGGATGGTGAAGTCTCTGCTGCTGCTGAGAGCTTC 358
34 InsertAlaProGlyProValLeuValProThrArgLeuGlnThrLeu 50
359 AGCTGCCCCAGGCGCTGTGCTAGTGCCTACCGCTCAGAGAGAGCTG 408
51 ValLeuArgCysProGlnLysThrAspCysAlaLeuArgValArgVal 67
409 GTCTGAGAGTCTCCACAGACAGATTGGCCCTCTGTCTCGGTGGT 458
67 ValAlHisLeuAlaValHisGlyHisTrpAlaGluProGluGlnAlaGly 84
459 GGTCACACTGGCCGTCATGGGACACTGGGAGAGCTGAGAGAGCTGGA 508
84 ySerAspSerGluLeuGlnGluSerArgAsnAlaSerLeuGlnAlaGln 100
509 AGCTGATTTCAGAACTCCAGAGAGCTGAGAGAGCTCTCTCCAGGCCAG 558
101 ValValLeuSerPheGlnAlaTrpProIleAlaArgCysAlaLeuLeuG 117
559 GTGGTGTCTCTCTCCAGGCTTACCCATCCGCTGCTGCTGCTGCTGA 608
117 uValGlnValProAlaAspLeuValGlnProGlyGlnSerValGlySer 134
609 GGTCCAGAGTCCCGCTGACCTGCTGAGCAGCTGCTGAGTCCGCTGCTG 658
134 ValAlHisLeuAspCysPheGlnAlaSerLeuGlnAlaGlnValGlnTrp 150
659 CGGTATTTCAGTCTTTCAGAGCTCTGAGGCTGAGTACAGATCTGG 708
151 SerTrpThrLysProArgTrpGlnLysGlnLeuAsnLeuThrGlnGln 167
709 TCTTACACGAAAGCCAGAGTACAGAAAGAGCTCAACTCAGACAGAGCT 758
167 uPro..... 168
759 GCCTGACTGACGAGGCTTGAAGTCCGGAGACAGATCCAGAGCTGCTGG 808
169 .....AspGlyAspAsnValLeuLeuThr 176

```

```

809 TCTTCCCTGGCTCAATGTGTCTACAGATGTGACATGTCTTCTGACA 858
177 LeuAspValSerGluGlnGlnAspPheSerPheLeuLeuTrpLeuArgPr 193
859 CTGATGTCTCTGAGAGAGCAGACCTTACCTTCTTACTGACTGGCTCC 908
193 oValProAspAlaLeuLysSerLeuTrpTrpLysAsnLeuThrGlyPro 210
909 AGTCCCGGATGCTCAAAATCTGTGTGATACAAAACCTGACAGACTTC 958
210 InAsnIleThrLeuAsnHisThrAspLeuValProCysLeuCysIleGln 226
959 AGAATATTACTTTAAACACACAGACTGTTCTCCCTGCTCATTCAG 1008
227 ValTrpSerLeuGluProAspSerGluArgValGluPheCysProPheAr 243
1009 GTGTGTGCTAGAGCCAGACTCTGAGAGGTCGAATTTCTGCCCCCTCCG 1058
243 gCysAspProGlyAlaHisArgAsnLeuTrpHisIleAlaArgLeuArgV 260
1059 GGAATATCCCGTGCACACAGACACTTGGCAGATAGCCAGGCTGCGGG 1108
260 aLeuSerProGlyValTrpGlnLeuAspAlaProCysCysLeuProGly 276
1109 TACTTCCCCAGGCGTATGCGAGCTGATGCGCTGCTGTGCGCGGGC 1158
277 LysValThrLeuCysTrpGlnAlaProAspGlnSerProCysGlnProle 293
1159 AAGTAAACACTGTGCTGGCAGGACACACAGACAGAGTCCCTCCAGGCCACT 1208
293 uValProProValProGlnLysAsnAlaThrValAsnGluProGlnAsp 310
1209 TGTGTCACACAGGCCCCAGAAAGCCACTGTGAATGAGCCCAAGAT 1258
310 heGlnLeuValAlaGlyHisProAsnLeuCysValGlnValSerThrTrp 326
1259 TCCAGTTGTGTGGCAGGCGCACCCCAACCTGTGTGTCCAGGTGACACTGG 1308
327 GluLysValGlnLeuGlnAlaLysSerTrpAlaAspSerLeuGlyProPh 343
1309 GAGAAAGTTCAGCTCAAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1358
343 eLysAspAspMetLeuLeuValGluMetLysThrGlyLeuAsnAsnThrS 360
1359 CAAGATATATATGTCTGTAGTGAAGTGAAGAAACCGGCTTCACACACAT 1408
360 erValCysAlaLeuGluProSerGlyCysThrProLeuProSerMetAla 376
1409 CAGTCTGTGCTTGAACCCAGCTGGGTGTACACACTGCCCAGACATGGCC 1458
377 SerThrArgAlaAlaArgLeuGlyGlnGluLeuLeuGlnAspPheArgSe 393
1459 TCCAGAGAGCTGTCTGCTGCTGGAGAGAGTGTCTGCAAGACTTCCGATC 1508
393 rHisGlnCysMetGlnLeuTrpAsnAspAspAspMetGlySerLeuTrp 410
1509 ACACCAAGTGTATGACCTGTGTGAACAGATGACACAGTGGATTCCTATGGG 1558
410 LaCysProMetAspLysTrpIleHisArgArg 420
1559 CCTGGCCCATGGACAAATACATCCACAGCGC 1590

```

seq_name: gb_pat:AX350978

seq_documentation_block:

LOCUS AX350978 2094 bp DNA linear PAT 06-FEB-2002
 DEFINITION Sequence 12 from Patent M00190356.
 ACCESSION AX350978
 VERSION AX350978.1 GI:18616354
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified
 unclassified.


```

REFERENCE 1 (sites)
AUTHORS Gorman,D.M.
TITLE Mammalian receptor proteins: related reagents and methods
JOURNAL Patent: WO 0190358-A 12 29-NOV-2001;
SCIERING CORPORATION (US)
FEATURES location/Qualifiers
SOURCE 1. 2094
/organism="unidentified"
/db_xref="taxon:32644"
/note="rodent; surmised Mus musculus"
BASE COUNT 247 a 251 c 393 g 285 t 918 others
ORIGIN

Alignment_scores:
Quality: 1729.00 Length: 441
Ratio: 5.056 Gaps: 1
Percent Similarity: 77.551 Percent Identity: 70.975

Alignment block:
US-09-899-471-2_COPY_21_440 x AX350978 ..

Align seg 1/1 to: AX350978 from: 1 to: 2094

2 GUATAGLeuMetGluProGlnAlaSerPThrAlaArgCysSerLeuGlyLeuSe 18
||| ||||||||||||||||||| |||::: ||| ::
64 GARGMGNTNATGAGACCCAGCAAGAAACGCMGNTGMSNTTNGCNGNTYMS 113
18 TCySHISLeuTPAspGlyAspValLeuCySLeuProGlySerLeuGlns 35
:::||||| |||||||||||| || |||||::: |||:
114 NTGYCATYTTNTGGAGTGGNGAGTGTNTGTCTYTTCCGCMGMSNTTNCARW 163
35 eRaLaProGlyProValLeuValProThrArgLeuGlnThrGluLeuVal 51
:::||||||||| ||||||| ||||||| |||
164 SNGCCMGCCGNCNGTNTTNGTCCNACMGNTYTCARACNGARNTNGT 213
52 LeuArgCySProGlnLysThAsPCysAlaLeuArgValAlrValValVa 68
||||||||| ||||||| ||| |||||||
214 YTNMGNTGYCCNCAAAACAGATYGCGTNTGTCTNGMNGTNGTNGT 263
68 IHSLeuAlaValHISGlyHISThrPAlaGluProGluGluAlaGlyLys 85
||| ||||||||||| ||||| ||||| |||||
264 NCAYTYTNCNGTNCATGCGNCATYTGCGCGNARCCNGARGARGCGNAAAW 313
85 eRaSPSerGluLeuGlnLysSerArgAsnAlaSerLeuGlnAlaGlnVal 101
:::||||| |||||::: |||||::: |||||
314 SNGAWMSNGARYTNCARGARMSNMGNAAYCGMSNTYTCARGCCNARGTN 363
102 ValLeuSerPheGlnAlaIaTYProIleAlaArgCysAlaLeuLeuGluVa 118
||| :::||||||||| ||||| |||||
364 GNTYNTSNTTYCARGCCNTATCCNATGCMGNTGCGCTNTNTTNCARGT 413
118 IGlInValProAlaAspLeuValGlnProGlyGlnSerValGlySerAlay 135
414 NCARGTNCNGCNCATYTTNGTNCARCCNGGNCARMSGTGGMMSGCMG 463
135 aLpHeaSPCySPhaeGluAlaSerLeuGlyAlaGluValGlnIleTPSer 151
||||||||| |||||::: |||||::: |||||:::
464 TMTTYGATGYTTCARGCMMSNTTNGCNGCNGARGTNCARATHTGMSN 513
152 TyrThrLysProArgTYrGlnLysGluLeuAsnLeuThrGlnLeuP 168
||||||||| ||||| ||| ||||||| ||
514 TATACNAARCCNMGNTATYCAAAAGARGYTNAAYTTCNARCARARNTCC 563
168 O..... 168
1 NGATYGMGNGSNTTNCARGTNNCGNAGTYSNATHCARMSVTGTGGCTTY 613
564
169 ..... AspLysAspAsnValLeuLeuThrLeu 177
614 TNCCTGGTYTNAAGCTMMSNACNGAYCNGATYAAYGNTNTYTNACNTYN 663
178 AspValSerGluGluGlnAspPheSerPheLeuLeuTyTLeuArgProVa 194

```

[illegible]

TITLE Human cytokine receptor
JOURNAL Patent: WO 0104304-A 1 18-JAN-2001;
ZymoGenetics, Inc. (US)

FEATURES
source location/qualifiers
1. 2255
/organism="Homo sapiens"
/db_xref="taxon:9606"
154. 2232
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28435.1"
/db_xref="GI:12710226"
/translation="MPVWFLSLALGRSPVLSLERLVGPQDATHCSPGLSCRLMDS
DILCPDIPAPVPVLAFTLQTELVLCOKETCDLCRAVLAHVGHMEPEDE
EKFGAADSVEEERNASLOAVLSPVAPARVILEVPAALNORGOSGVSVY
DCEPALDSERVITSYORPREKELNTOOLPALPMLVSAIDNVHVLANSBOHF
GSLYRNQVGGPRPRMKNLIGPQITLNTDLPCLQVWPLEPDSVPTNICEPR
EDRAHONMQAAARLLTLQSWLDAPCSLPAEALMCRWAPGDCPLVPLSMEN
VYDVKLEPFLKGNPCVQVNSSEKIQEQLQADSLDPLKDDVLLLETGPODNR
SLCALBESGCTSLPKASTRAARAGEYLQDLSGQCLQIMDDDLGALMACPMKXTH
KRWALWLAACLFPAALSLIILKKDKAKAARGAAILYSADSGPERIVGALASA
LCOLPLRAVDIHSRRLSNQPYANFHAQRQTIQEGCVVLLSPGVALVCEWLQ
DVGSGRGAHPHAFSLSCVLPDPLAGRAGSIVGACFPLLPDAVPALEFVIV
FTLPQDPFLGALQDPRAPRSRLQERAEVSRALQPLDSYFHPPTPAPGRGVGP
GAPGAGDGT"

CDS
154. 2232
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28435.1"
/db_xref="GI:12710226"
/translation="MPVWFLSLALGRSPVLSLERLVGPQDATHCSPGLSCRLMDS
DILCPDIPAPVPVLAFTLQTELVLCOKETCDLCRAVLAHVGHMEPEDE
EKFGAADSVEEERNASLOAVLSPVAPARVILEVPAALNORGOSGVSVY
DCEPALDSERVITSYORPREKELNTOOLPALPMLVSAIDNVHVLANSBOHF
GSLYRNQVGGPRPRMKNLIGPQITLNTDLPCLQVWPLEPDSVPTNICEPR
EDRAHONMQAAARLLTLQSWLDAPCSLPAEALMCRWAPGDCPLVPLSMEN
VYDVKLEPFLKGNPCVQVNSSEKIQEQLQADSLDPLKDDVLLLETGPODNR
SLCALBESGCTSLPKASTRAARAGEYLQDLSGQCLQIMDDDLGALMACPMKXTH
KRWALWLAACLFPAALSLIILKKDKAKAARGAAILYSADSGPERIVGALASA
LCOLPLRAVDIHSRRLSNQPYANFHAQRQTIQEGCVVLLSPGVALVCEWLQ
DVGSGRGAHPHAFSLSCVLPDPLAGRAGSIVGACFPLLPDAVPALEFVIV
FTLPQDPFLGALQDPRAPRSRLQERAEVSRALQPLDSYFHPPTPAPGRGVGP
GAPGAGDGT"

BASE COUNT 369 a 748 c 703 g 435 t
ORIGIN

alignment_scores:
Quality: 1554.00 Length: 433
Ratio: 4.281 Gaps: 3
Percent Similarity: 83.834 Percent Identity: 67.667

Alignment block:
US-09-899-471-2_COPY_21_440 x AX073993 ..

Align seq 1/1 to: AX073993 from: 1 to: 2255

```

1 LeuGluArgLeuMetGluProGlnAspThrAlaAspCysSerLeuGlyLe 17
|||||
214 CTGAGAGAGCTTGAGGCTCTGAGACCTACCTGCTCCGGGCT 263
17 userCysHisLeuTrpAspGlyAspValLeuCysLeuProGlySerLeu 34
CTCTGCGCGCTCTGGGACAGACATACCTGCTGCTGGGACATCG 313
34 InsertAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeu 50
|||||
314 TGCTGCTCCGGGCGCCCTGCTGGGCGCTACGACCTCGACAGAGCTG 363
51 ValLeuArgCysProGlnLysThrAspCysAlaLeuArgValAlaVal 67
|||||
364 GTGCTGAGGTCCAGAGAGACACGACTGCTGCTGCTGCTGCTGCTG 413
67 ValHisLeuAlaValHisGlyHisThrAlaGlnProGlnGlu..... 81
|||||
414 TGTCACCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 463
82 .....AlaGlyLysSerAspSerGluLeuGlnGluSerArgAsnAlaSer 96
AGTTTGAGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 513
97 LeuGlnAlaGlnValValLeuSerPheGlnAlaIleProIleAlaArgC 113
|||||
514 CTCACAGGCCAAGTGTGCTCTCTCCAGAGCTACCTACTGCGCGCTG 563
113 salaleuLeuGlnValGlnValProAlaAspLeuValGlnProGlyGln 130
|||||
564 CGTCTCTGTGAGGTGCAAGTGCCTGCTGCTGCTGCTGCTGCTGCTG 613
130 erValGlySerAlaValAlaPheAspCysPheGlnAlaSerLeuGlnAla 146
|||||

```

```

614 CTGTGGGCTGTGTGATATGACTGTTGAGAGCTGCCCTAGGAGTGAG 663
147 ValGlnIleTrpSerTrpThrLysProArgTrpGlnLysGluLeuAsn 163
|||||
664 GTAGAGATCTGTGTCTATCTACGACCCAGTAGAGAGAACTCAACCA 713
163 uThrGlnGlnLeuPro.....AspGlyA 171
|||||
714 CACACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 763
171 spAsnValLeuLeuThrLeuAspValSerGluGlnAspPheSerPhe 187
|||||
764 ACAAGCTGATCTGTTCTGAATGTCTCTGAGAGACCACTTGCGCTC 813
188 LeuLeuTrpLeuAlaProValProAspAlaLeuLysSerLeuTrpTr 204
|||||
814 TCCTCTACTGGAATCAGTCCAGGGCCCCCAAAACCCCGTGCACAA 863
204 sAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAspLeuAl 221
|||||
864 AAACCTGACTGACCGCAGATCATCTTACCTGAACACACACACCTGTC 913
221 rocYsLeuGlyIleGlnValTrpSerLeuGlnProAspSerGlnArg 237
|||||
914 CCTGCTGTATTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963
238 GluPheCysProPheArgGluAspProGlyAlaHisArgAsnLeuTr 254
|||||
964 AACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 1013
254 sIleAlaArgLeuArgValLeuSerProGlyValTrpGlnLeuAspAl 271
|||||
1014 ACCGCGCCGAGCTGAGCTGCTGACCTGACCTGACAGCTGCTGCTG 1063
271 rocYsCysLeuProGlyLysValThrLeuCysTrpGlnAlaProAsp 287
|||||
1064 CGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1113
288 SerProCysGlnProLeuValProProValProGlnLysAsnAlaThr 304
|||||
1114 GACCCCTGCGCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1163
304 LAsnGluProGlnAspPheGlnLeuValAlaGlyHisProAsnLeu 321
|||||
1164 GCACAGGCTTCTGAGTTCATGCTGTAAGAGCCACCTTAACCTGTG 1213
321 aGlnValSerThrTrpGluValGlnLeuGlnAlaCysSerTrpAla 337
|||||
1214 TTCAGTGAACAGCTGCGAGAACCTGACGCTGACGAGCTGTTGTGCT 1263
338 AspSerLeuGlyProPheLysAspAspMetLeuValGluMetLysThr 354
|||||
1264 GACTCCCTGCGGCTCTCAAAAGAGATGCTACTGTTGAGACAGAGG 1313
354 rGlyLeuAsnAsnThrSerValCysAlaLeuGlnProSerGlySerThr 371
|||||
1314 CCCCAGAGCAACAGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1363
371 rGluProSerMetAlaSerThrArgAlaIleArgLeuGlyGluGluLeu 387
|||||
1364 CACTACCCAGCAAGCTCTCCAGAGGCGACCTGCTGCTGCTGCTGCT 1413
388 LeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsnAspAs 404
|||||
1414 CTACAGACCTGACAGTCCAGGCGCAGTCTCTGCTGCTGCTGCTGCT 1460
404 nMetGlySerLeuTrpAlaCysProMetAlaPylsTyrlHisArgArg 420
|||||
1461 CTGGGAGCGCTATGGGCTGCGCCCATGAGCAATATCATCCACAAG 1509
seq_name: gb_pat:AX092430
seq_documentation_block:

```


LOCUS AX092430 2380 bp DNA linear PAT 21-MAR-2001
 DEFINITION Sequence 161 from Patent WO0116318.
 ACCESSION AX092430
 VERSION AX092430.1 GI:13444529
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 2380)
 Eaton, D.L., Filvaroff, E., Gerritsen, M.E., Goddard, A.,
 Godowski, P.J., Grimaldi, C.J., Gurney, A.L., Watanabe, C.K. and
 Wood, W.I.
 TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
 the same
 JOURNAL Patent: WO 0116318-A 161 08-MAR-2001;
 Genentech, Inc. (US)
 FEATURES
 source 1. 2380
 Location/Qualifiers
 BASE COUNT 411 a 776 c 743 g 450 t
 ORIGIN
 Alignment_scores:
 Quality: 1554.00 Length: 433
 Ratio: 4.281 Gaps: 3
 Percent Similarity: 83.834 Percent Identity: 67.667
 Alignment_block:
 US-09-899-471-2_COPY_21_440 x AX092430 ..
 Align seg 1/1 to: AX092430 from: 1 to: 2380

743 GTACGATCTGTCTTACTACACCCAGGTACGAGAGAACTCAACCA 792
 163 urthglnlnleupro.....asp1ya 171
 793 CACACAGACAGCTGCTCCCTGCTGCTCAGCTGCTCAGACAGATGCTG 842
 171 spasnvalleuileuaspvalsergluuglnasptheserthe 187
 843 ACACGTGATCTGTGTCTTAATCTCTGTGAGAGACACTTGGCTG 892
 188 leuileutyrleuargprovalproaspalauleuysserleutpytly 204
 893 TCCCTGTACTGAGATCAGTCCAGAGGCCCCCAAAACCCGCTGACACA 942
 204 spasnleuthgilyproglinsnillethleuasnhisrthaspleuvalp 221
 943 AAACCTGACTGAGACCCGAGATCATTTACTTGAACACACACACCTGCTC 992
 221 rocysleucyslleglnvaltrpserleuugluproaspsergluargval 237
 993 CCGCTCTGTATTCAGGTGTGCTGCTGGAACCTGACTCGCTGTAAGACG 1042
 238 gluupheCysPropheargluaspproglialahisargasnleutrpH 254
 1043 AACATCTGCCCTTCAGAGAGACCCCGCAGACACCAACCTCTGSCA 1092
 254 silalaargleuargvalleuserproglialtrp6lnleuaspalap 271
 1093 AGCCGCCGCGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 1142
 271 rocysCysleuproglylylvalthrleucystyrp6lnalaproasp6ln 287
 1143 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1192
 288 serproCysgluProleuvalproprovalproglinsasnalaathva 304
 1193 GACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1242
 304 lasn6lnuproglnaspheglnleuvalalaglYHISProbsnleucysv 321
 1243 GGACAAAGTTCTGAGTTCCATTGCTGGAAGGCAACCACTGCTGCTG 1292
 321 a6lnvalserthrtrp6lnlylval6lnleu6lnalacysSerrtpala 337
 1293 TTCAGGTGAACAGCTGCTGCACTGCACTGCACTGCACTGCTGCTGCTG 1342
 338 aspserleu6lnProphelyasaspmetleuval6lnumetlysth 354
 1343 GACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1392
 354 r6lyleuasnasntrhserValCysalaleu6lnuproserr6lyCysThrP 371
 1393 CCCCAGACAGACAGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1442
 371 roleupProsermetalaSerthrArAlaAlaarg6ln6lyglu6lnleu 387
 1443 CACTACCCCAAGAGGCTGCTGCACTGCACTGCACTGCTGCTGCTGCTGCT 1492
 388 leu6lnaspheargserth6ln6ln6ln6ln6ln6ln6ln6ln6ln6ln6ln 404
 1493 CTACAAAGACTGCACTGCACTGCACTGCACTGCTGCTGCTGCTGCTGCTG 1599
 404 m6et6lyserleu6lnalacysPrometasplystYrllHISargarg 420
 1540 CTGGGAGAGGCTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1588
 seq_name: gb_pat:AX180776
 seq_documentation_block:
 LOCUS AX180776 2380 bp DNA linear PAT 06-AUG-2001
 DEFINITION Sequence 13 from Patent WO0146420.
 ACCESSION AX180776
 VERSION AX180776.1 GI:15132622

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 2380)

Chen, J., Filvaroff, E., Fong, S., Goddard, A., Godowski, P. J., Grimaldi, C. J., Gunney, A. L., Li, H., Hillan, K. J., Tumas, D., van Lookeren, M., Vandlen, R. L., Watanabe, C. K., Williams, P. M., Wood, W. I. and Yansura, D. G.

11-17 and 11-17r homologous polypeptides and therapeutic uses thereof

Patent: WO 0146420-A 13 28-JUN-2001.

FEATURES

Location/Qualifiers

BASE COUNT 411 a 776 c 743 g 450 t

ORIGIN

alignment_scores:

Quality: 1554.00 Length: 433
Ratio: 4.281 Gaps: 3
Percent Similarity: 83.834 Percent Identity: 67.667

alignment_block:

US-09-899-471-2_COPY_21_440 x AX180776

Align seg 1/1 to: AX180776 from: 1 to: 2380

1 LeuGluuArgLeuMetGluProGlnAspThrAlaArgCysSerLeuGlyLe 17
 293 CTGGAGAGGCTGTGGGGCCCTCAGAGACCTACCCCTGCTCCGGGCT 342
 17 uSerCysHisLeuTrpAspGlyAspValLeuCysLeuProGlySerLeu 34
 343 CTCCTGGCCCTCTGGGACAGTACTCTGCTGCTGCTGGGACATCG 392
 34 InseAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeu 50
 393 TGGCTGCTCCGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 442
 51 ValLeuArgCysProGlnLysThrAspCysAlaLeuArgValArgVala 67
 443 GTGCTGAGTGCCAGAGAGACCGACGTGACCTGCTGCTGCTGCTG 492
 67 ValHisLeuAlaValHisGlyHisTrpAlaGluProGluGlu..... 81
 493 TGTCCACTTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 542
 82AlaGlyLysSerAspSerGluLeuGlnLysArgAsnAlaSer 96
 543 AGTTTGGAGGACGACGTACTCAGAGGGTGGAGGACCTAGATCCCT 592
 97 LeuGlnAlaGlnValValLeuSerPheGlnAlaTyrProIleAlaArgC 113
 593 CTCAGAGCCCAAGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 642
 113 AlaAlaLeuGluValGlnValProAlaAspLeuValGlnProGlyGln 130
 643 CCTGCTGCTGAGAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTG 692
 130 eValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaG 146
 693 CTGTGGGCTGTGTGTATATACGCTTCTGAGGCTGCTGAGGAGTGA 742
 147 ValGlnIleTrpSerTyrThrLysProArgTyrGlnLysGluLeuAsn 163
 743 GTACGATCTGTCTTACTACAGCCCAAGTACGAGAGGAGCACTCA 792
 163 uThrGlnGlnLeuPro.....AspGlyA 171

793 CACACAGCAGCTGCTGCCCTGCTGCTCAGCGTGTACAGAGATGTGT 842
 171 spAsValLeuLeuThrLeuAspValSerGluGlnAspPheSerPhe 187
 843 ACACGTCATCTGTCTGTAAGTCTCTGAGAGACGACCTCGGCTC 892
 188 LeuLeuTrpLeuArgProValProAspAlaLeuLysSerLeuTrpTyr 204
 893 TCCCTGTACTGATACAGTCCAGGCCCCCAAAACCCCGGTGGACA 942
 204 sAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAspLeuVal 221
 943 AAACCTGACTGACCGCAGATCATATACCTGAAACACACAGACTGTT 992
 221 rOCysLeuCysIleGlnValITrPserLeuGlnProAspSerGluArg 237
 993 CTTGCTGTGTATTCAGGTGGCTCTGGAACCTGACTCTGTTAGGAG 1042
 238 GluPheCysProPheArgGluAspProGlyAlaHisArgAsnLeuTrp 254
 1043 AACATCTGCCCTTCAGAGAGACCCCGCGCACACACAGAACTGTGCA 1092
 254 sIleAlaArgLeuArgValLeuSerProGlyValITrPLeuAspAla 271
 1093 AGCCCGCCGACTGCGACTGCTGACCTGCAGAGACTGCTGCTGAGCC 1142
 271 rOCysCysLeuProGlyLysValIThrLeuCysITrPLeuAlaProAsp 287
 1143 GTGCTGCTGCTGCCCGCAGAGGCGCAGTGTGCGGGCTCGGGTGG 1192
 288 SerProCysGlnProLeuValProProValProGlnLysAsnAlaTr 304
 1193 GACCCCTGCCACCACTGTCTCCACCGCTTCTCTGGAGAACCTGCTGT 1242
 304 lAsnGluProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuC 321
 1243 GGACAAAGTTCGAGTTCATTTGCTGAAGGCCAACCTGATCTGTG 1292
 321 alGlnValSerThrITrPLeuLysValGlnLeuGlnAlaCysSerITr 337
 1293 TTCAGTGAACAGCTCGGAGAGCTGCTGCTGCTGCTGCTGCTGCT 1342
 338 AspSerLeuGlyProPheLysAspMetLeuLeuValGluLeuLys 354
 1343 GACTGCTGCGGCTCTCAAAAGCATGTGCTACTGTGGAGACAGGAG 1392
 354 rGlyLeuAsnAsnITrSerValCysAlaLeuGluProSerGlyCysThr 371
 1393 CCCCAGAGCAACAGATCCCTGCTGTGCTGCTGCTGCTGCTGCTG 1442
 371 rOleuProSerMetAlaSerITrArgAlaAlaITrLeuGlyGluLeu 387
 1443 CACTACCCAGCAAAAGCTCCAGAGGCGAGCTGCTGAGAGTACT 1492
 388 LeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsnAsp 404
 1493 CTACAGACCTGCACTACAGCCAGGTGTCTGCAAGTGG...GACGATGA 1539
 404 mMetGlySerLeuTrpAlaCysProMetAspLysTyrIleHisArg 420
 1540 CTTGGGAGCGCATGTGGCTGCGCCATGAGCAAAATCATCATCAAGGC 1588

seq_name: gb-BC006411
 seq_documentation_block:
 LOCUS BC006411 2507 bp mRNA linear PRI 12-JUL-2001
 DEFINITION Homo sapiens, clone MGC:10763 IMAGE:3606276, mRNA, complete cds.
 ACCESSION BC006411
 VERSION BC006411.1 GI:13623590
 KEYWORDS MGC;
 SOURCE human.
 ORGANISM Homo sapiens

258 CTGGAGAGGCTTGTGGGGGCTTCAGAGACGCTAACCCACTGCTCTCCGGGGCT 307

17 uSerCysHisLeuTrpAspGlyAspValLeuCysLeuProGlySerLeu 34

308 TTTCTTCCCGCCCTCTGTGGACAGTGAATCTCTGCTGCTGGGGACATCG 357

34 InSerAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeu 50

358 TCCCTGCTCCGGGGCCCGTGTGGGGCTTACGCACTCGACACAGACGTG 407

51 ValLeuArgCysProGlnLysThrAspCysAlaLeuArgValAlaVal 67

408 GTGCTGAGGTGTCACAGAGAGACGACATGTGACCTGTGCTGCTGGTGGC 457

67 ValHisLeuAlaValHisGlyHisThrAlaGluProGluGlu..... 81

458 TGTCCACTGTGGCGGTGCATGGGCACTGGGAACAGCTGAAGTGAAGAA 507

82AlaGlySerAspSerGluLeuGlnGluSerArgAsnAlaSer 96

508 AGTTGGAGAGACACACTGACTGATCGGGGTGGAGAGACCAAGAAATCCT 557

97 LeuGlnAlaGlnValValLeuSerPheGlnAlaTyrProIleAlaArgC 113

558 CTCACGGCCCAAGTGTGCTCTCTTCCAGGCGTCACTACGCCCCGCTG 607

113 SAlaLeuLeuGluValGlnValProAlaAspLeuAlaGlnProGlyGln 130

608 CTTCTGTGGGAGGTGCAGATGCTCGTGTGCTTGTGAGTTGTGTCACT 657

130 erValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGlu 146

658 CTGTGGGCTGTGGTTATATAGCTCTTCGAGGCTGCCCTTAGGAGTAG 707

147 ValGlnIleTrpSerTyrThrLysProArgTyrGlnLysGluLeuAsnLe 163

708 GTACGAATCTGCTCTTAATACAGCCAGATGACGAAGAACTCAACCA 757

163 uThrGlnGlnLeuPro.....AspGlyA 171

758 CACACAGACAGCTGCTGCCCTGCCCTGACTCAAGCTGCACAGATGGTG 807

171 spAsnValLeuLeuThrLeuAspValSerGluGlnGlnAspPheSerPhe 187

808 ACAACGTGCTATCTGTTCGATATGTCTGTGAGGACACACTTCGGCTTC 857

188 LeuLeuTyrLeuArgProValProAspAlaLeuLysSerLeuTyrTyr 204

858 TCCCTGTACTGGAATCAGGTCCAGGGCCCCCAAAACCCGGTGCACAA 907

204 sAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAspLeuVal 221

908 AAACCTGACTGACGCGCAGATCATTAACCTTGAAACACACAGCTGGTTC 957

221 roCysLeuCysIleGlnValTrpSerLeuGluProAspSerGlnArgVal 237

958 CTGCTCTGTATTTAGGGTGGCTCTGGAAACCTGACTCCGTAAGAGACG 1007

238 GluPheCysProPheArgGluAspProGlyAlaHisAspAsnLeuTrpHis 254

1008 AACATCTGCGCTTCAGGAGAGACCCCGCGACACCAACAACTGTGGCA 1057

254 sIleAlaIArgLeuArgValLeuSerProGlyAlaTrpGlnLeuAspAla 271

1058 AGCCCGCCGACGACGACATGCTACCTCGACAGAGCTGGCTGTGACGAC 1107

271 roCysCysLeuProGlyLysValThrLeuCysTrpGlnAlaProAspGln 287

1108 CTGTCTGCTGCTGGCCGACAAAGCGGACATCTGCTGGCGGGCTCCGGGTGG 1157

288 SerProCysGlnProLeuValProProValProGlnLysAsnAlaThrVa 304

1158 GACCCCTTCGACGCCACTGTCTCCACCGGCTCTCTGGGAAAGAGCTCATCTGT 1207

304 LAsnGluProGlnAspPheGlnLeuValAlaGlyHisProAsnLeucysv 321
 1208 GGACAGAGTTCTCGAGTTCCTCCATGCTCAAGAGCCACCCCTCTGCTG 1257
 321 AlGlnValSerThrTrpGlnValGlnLeuGlnAlaCysSerTrpAla 337
 1258 TTCAGGTGAACAGCTCGGAGAGCTCCAGCTGAGAGTCTTGAGGCT 1307
 338 AsperLeuGlyProPheLysAspPheLeuValGlnMetLysTh 354
 1308 GACTCCCTGGGGCTCTCAAGACGATGTGCTACTGTGGAGACAGAGG 1357
 354 rGlyLeuAsnThrSerValCysAlaLeuGluProSerGlyCysThr 371
 1358 CCCCCAGACACAGATCCCTCTGCTTGGACACAGCTGCTACTT 1407
 371 rGluProSerMetAlaSerThrArgAlaAlaArgLeuGlyGlnLeu 387
 1408 CACTACCCAGCAAGCCTCCAGAGGAGCTGCTTGGAGAGTCTTA 1457
 388 LeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsnAspAsp 404
 1458 CTACAGACCTGACGTGAGGAGGAGCTGTGAGCTATGG...GACATGA 1504
 404 nMetGlySerLeuTrpAlaCysProMetAspLysTrpLysArgArg 420
 1505 CTGGGAGGCTGATGGCTGCTCCATGACAAATACATCCACAGAGCCG 1553
 seq.name: gb_pat:AX350973

seq documentation block:
 LOCUS AX350973 2308 bp DNA linear PAT 06-FEB-2002
 DEFINITION Sequence 7 from Patent W00190358.
 ACCESSION AX350973
 VERSION AX350973.1 GI:18616349

SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source

unidentified.
 unidentified.
 1 (sites)
 Gorman, D.M.
 Mammalian receptor proteins: related reagents and methods
 Patent: WO 0190358-A 7 29-NOV-2001;
 SCHERING CORPORATION (US)
 Location/Qualifiers
 1. 2308

CDS

/organism="unidentified"
 /db_xref="taxon:32644"
 /note="primate; surmised Homo sapiens"
 181..2292
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD22835.1"
 /db_xref="GI:18616350"
 /translation="MPVWFLLSLALGRSOMILSLERLVSPQDATHTCSPLGSLRLMDS
 DILCEPGDIPAPGPAVLAFTHLQTELVLCOKETDCDLCRLVAHLVHGVHWPED
 KEFGAAGDVEEPNNAALQAVLSFOAYPTARCVLLEOVPAVALVOFGSVSVY
 DCFEALGSEVRIMSTQDPRYERELNHTQQLPDCRGLEVNNSITSCALPRLANSADG
 DNVIHLVANSSEOHRLSLXMYNOVCPKPRKRNKLTGPOITITNHTDLVPLCIQVW
 BLEPVSRYNTICPFREDPRRHOHLWQARLRLITLWSLWDAPSLPAEALCWRAFG
 GPCQPLVPELSTMENTVDVNSSEKLQLOECIMWADSLPKDKDVLLETRGPONRSL
 CALBESGCTSLPASKASTRAARLGEVLILOLQSGCCQLMDDCALACPMKDYIKR
 MALVIALCLFPAALSLILLLKDKHAKGMLRLIKODVRSAAARGRALILLYSDDSG
 FERLVGALASALCOLPLRVAVDLMSRRELISAQPVMEHQROROTLOEGCYVILFSP
 GAVATCSEMLQDGVSGRAHGRHDAKRAISCTVLPDTLQRAPESTYVACGCDRLHPD
 AVPALEFRVPTVTLTSLQDLDFCALQDQPRAPRSGRLOERAQVSRALQALPDLSYFHP
 GTSAPGRGVGPGAGPGAGGT"
 241..2289
 mat.peptide
 /product="unnamed"

BASE COUNT 384 a 757 c 731 g 434 t 2 others
 ORIGIN

alignment scores:
 Quality: 1465.00 Length: 448
 Ratio: 4.198 Gaps: 4
 Percent Similarity: 77.902 Percent Identity: 62.946

alignment block:
 US-09-899-471-2_COPY_21_440 x AX350973

Align seg 1/1 to: AX350973 from: 1 to: 2308

1 LeuGlnAlaLeuMetGluProGlnAspThrAlaArgCysSerLeuGly 17
 241 CTTGAGAGGCTTGTGGGCTCAGAGAGCTACACCACTGCTCCGGGCT 290
 17 uSerCysHisLeuTrpAspGlyAspValLeuGlyLeuProGlySerLeu 34
 291 CTCTGCTCCGCTCTGGGAGTACATATCTGCTGCTGCTGGGAGATCG 340
 34 InSerAlaProGlyProValLeuValProThrArgLeuGlnThrGlnLeu 50
 341 TCCCTGCTCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 390
 51 ValLeuArgCysProGlnLysThrAspCysAlaLeuArgValArgVal 67
 391 GTCTGAGGTGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 440
 67 ValHisLeuAlaValHisGlyHisThrAlaGluProGlnLeu 81
 441 TGTCCACTGGCGCTGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAA 490
 82AlaGlyLysSerAspSerGluLeuGlnGlnSerArgAsnAlaSer 96
 491 ACTTGGAG 540
 97 LeuGlnAlaGlnValValLeuSerPheGlnAlaThrProIleAlaArgCys 113
 541 CTCAGAGCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 590
 113 SalLeuLeuGlnValGlnValProAlaAspLeuValGlnProGlyLys 130
 591 CGTCTGCTGGAGGTGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
 130 erValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGlu 146
 641 CTGGGCTGCTGTAATGATGATGATGATGATGATGATGATGATGATG 690
 147 ValGlnIleTrpSerTrpThrLysProArgTrpGlnLysGlnLeuAsnLe 163
 691 GTACAAATCTGCTCTATATCTACAGCCAGGTACGAGAGAGACTCAACA 740
 163 uThrGlnLeuPro 168
 741 CACACAGAGAGTGTCTACTGACAGGGGCTCGAAGTGTGAGACAGATCC 790
 169AspGlyAspAsn 172
 791 CGAGCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 173 ValLeuLeuThrLeuAspValSerGlnGlnAlaAspPheSerPheLeu 189
 841 GTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 890
 189 uTyrLeuArgProValProAspAlaLeuLysSerLeuTrpTyrLysAsn 206
 891 GTACTGATATGATGCTCAGAGGCCCCCAAAACCCCGGTGGACAAAAAC 940
 206 euThrGlyProGlnAsnIleThrLeuAsnHisThrAspLeuValProGys 222
 941 TGACTGAGAGCCGATATCTGATCTGAAACACACACAGACCTGCTCCCTG 990
 223 LeuCysIleGlnValTrpSerLeuGluProAspSerGluArgValGluPh 239

991 CTCTATATTCAGGTGTGGGCTCTGGAACTGACTGCTGCTTAGACGAACT 1040

239 eCysProPheArgGluAspProGlyAlaHisArgAsnLeuTrpHisIleA 256

1041 CTGGCCCTTCACGGGAGGAGCCCGCGGACACCAAGAACTCTGGCAAGCG 1090

256 IaArgLeuArgValLeuSerProGlyValTrpGlnLeuAspIleProCys 272

1091 CCGGACTGCGACTGCTGACCTGCGAGAGCTGCTGCTGGAGCCACCGTCC 1140

273 CysLeuProGlyLeuValThrLeuCysTrpGlnAlaProAspGlnSerP 289

1141 TCGGTGCCCCGAGAGGCGGACACTGTGTGGGGGGCGTCCGGGTGGAGACC 1190

289 OCysGlnProLeuValProProValProGlnLysAsnAlaTrpValAsnG 306

1191 CTGCAGCGCACAGTGTCCACCCCTTCTCGTGAGAGATGTCACTGTGAGC 1239

306 LuProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuCysValGln 322

1239 1239

323 ValSerThrTrpGluLysValGlnLeuGlnAlaCysSerTrpAlaAsp 339

1240 GTGACACAGCTCGAGACAGCTGACGTGCGAGAGTGTGTGGCTGTACTC 1289

339 rLeuGlyProPheLysAspAspMetLeuLeuValGluMetLysThrGlyL 356

1290 CCGTGGGGCTCTCAAGAGCATGTGTCTACTGTGGAGACAGCGGGCCCC 1339

356 euAsnAsnThrSerValCysAlaLeuGlnProSerGlyCysThrProLeu 372

1340 AGGACACAGAGCCCTGTGTGCTTGGAAACCAAGTGTGTACTTCACAT 1389

373 ProSerMetAlaSerThrArgAlaAlaArgLeuGlyGlnGlnLeuLeu 389

1390 CCCACACAAGCTCCACAGAGGCGACGTCCCTTGGAGATCTTACTTACA 1439

389 naSPheArgSerHisGlnCysMetGlnLeuTrpAsnAspAspMetG 406

1440 AGAGCTGACAGTCAGGCGCATGTGTCTGACCTATG... GAGATGACTTGG 1486

406 lySerLeuTrpAlaCysProMetAspLysTrpIleHisArgArg 420

1487 GAGCGCTATGGGCTGCCCCATGAGACAATAATCATCCACAAGCGC 1530

seq_name: gb_pat:AX136341

seqs documentation block:

LOCUS AX136341 2499 bp DNA linear PAT 30-MAY-2001

DEFINITION Sequence 263 from Patent EP1067182.

ACCESSION AX136341

VERSION AX136341.1 GI:14272747

KEYWORDS human.

SOURCE

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 2499)

Ota, T., Isogai, T., Nishikawa, T., Kawai, Y., Sugiyama, T. and Hayashi, K.

Secretory protein or membrane protein

Patent: EP 1067182-A 263 10-JAN-2001;

Helix Research Institute (JP)

location/Qualifiers

1..2499

/organism="Homo sapiens"

/db_xref="taxon:9606"

208..1011

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC39812.1"

/db_xref="GI:14272748"

FEATURES

source

CDS

```

BASE COUNT      418 a      821 c      764 g      496 t
ORIGIN

alignment_scores:
    Quality: 1397.00      Length: 504
    Ratio: 4.061      Gaps: 6
    Percent Similarity: 68.254      Percent Identity: 54.960

alignment_block:
US-09-899-471-2_COPY_21_440 x AX136341 ..
Align seg 1/1 to: AX136341 from: 1 to: 2499

1 LeuGIaArgLeuMetGlnProGlnAspThrAlaArgCysSerLeuGlyLe 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
268 CTGGAGAGAGCTGTGTGGGGCCCTCAGGAGCTACCCACTGCTCTCCGGGCTT 317
17 uSerCysHisLeuTrpAspGlyAspValLeuGlyLeuProGlySerLeuG 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
318 CTCCTCCGCCCTCTGGGACAGTGACTACTCTGCCCTGGGGGACATCG 367
34 InsertAlaProGlyProValLeuValProThrArgThrArgLeuIhrGluLe 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
368 TGCCTCTCTCCGGGCCCGCTGCTGGCCCTTACGCACTCGACAGAGAGCTG 417
51 ValLeuArgCysProGlnIhrAspThrAspCysAlaLeuArgValArgVala 67
418 GTGCTCAGGTGCCAGAGAGAGACCGACTGTGACCTGTCTGTGGTGTGGC 467
67 IValHisLeuAlaValHisGlyHisTrpAlaGluProGluGlu..... 81
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
468 TGTCACCTTGCCGCGTGCATGGGCACTGGGAGAGAGCGCTGAAGAGAGAAA 517
82 .....AlaGlySerAspSerGluLeuGlnGluSerArgAsnAlaSer 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
518 AGTTTGGAGAGCAGCTGACTCAGGGGCTGGAGAGAGCTTGAGAAATGCTCT 567
97 LeuGlnAlaGlnValValLeuSerPheGlnAlaTyrProIleAlaArgCys 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
568 CTCACAGGCCCAAGTCTGCTCTCTCCAGGCCCTACCTACTGCCCCGCTG 617
113 sAlaLeuLeuGlnValGlnValProAlaAspLeuValGlnProGlyGlns 130
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
618 CGTCCGTGGGAGGTGCAGAGTGCCTCTCCCTTGTGCAGTTGGTGCACAT 667
130 eValAlaIySerAlaValPheAspCysPheGlnAlaSerLeuGlnValGln 146
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
668 CTGTGGGCTCTGTGTATATGACTGCTTGGAGGCTGCTCAGGAGATGAG 717
147 ValGlnIleTrpSerTyrThrIysProArgTyrGlnValGluLeuAsnLe 163
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
718 GTACGAAATGTGGTCTATACTCAGGCCAGGTGCAGAGTACTCAACA 767
163 uThrGlnGlnLeuPro.....AspGlyA 171
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
768 CACAGGAGCTGCTGCGCCCTGCGCTCAACGTGACGAGAGATGAG 817
171 spAsnValLeuLeuThrLeuAspValSerGlnGlnGlnAspPheSerPhe 187
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
818 ACAACGTGATCTGTTGTGAATGTCTTGAGAGAGAGACTTGGGCTTC 867
188 LeuLeuTyrLeuArgProValProAspAlaLeuIySerLeuTrpTyrIy 204
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
868 TCCCTTACTGTGAATCAGTCCAGAGGCCCCCAAAACCCCGGTGGCACA 917
204 sAsn..... 205
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



```

918 AACCGGTGAGGCGTCCCGCTTCCCAAGTCATCCCATCTGTAGCCGA 967
205 .....
968 TCCCTGTGCAGAGGACGACATATCAGAGAGATCTTGAAGAGGA 1017
205 .....
1018 CTCACCCCAAGCAAGGAAATTTGGGGGAACTTCTCTCTGTT 1067
205 .....
1068 TCCCTTACTTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1117
206 .....
1118 CTTTATTTGTTCCACAGACTGAGCCGACATCATTAACCTGAACACAC 1167
217 TASPLeuValProCysLeuCysIleGlnValTTPSerLeuGluProAspS 234
1168 AACCTAGTTCCTGCTCTGTATTCAGTGTGGCTTGGAACTGACT 1217
234 GCGluArgValGluPheCysProPheArgGluAspProGlyAlaHisArg 250
1218 CCGTTAGGACGACATCTGCCCCCTCAGGGAGGACCCCGCGCACACG 1267
251 AsnLeuTTPHisIleAlaArgLeuArgValLeuProGlyValTTPG1 267
1268 AACCTTGGCAACCGCCCGACCTGCGACTGTACCCCGCAGAGCTGGCT 1317
267 nLeuAspAlaProCysCysLeuProGlyLysValThrLeuCysTTPGlna 284
1318 GCTGGACGACCGCTGCTGCTGCCCGCAGAAAGCGGACACTGTGGCGGG 1367
284 LaProAspGlnSerProCysGlnProLeuValProProValProGlnLys 300
1368 CTCGGGAGGAGACCCCTCCAGCCACACTGTGTCCACCGCTTCTGGAG 1417
301 AsnAlaThrValAsnGlnProGlnAspPheGlnLeuValAlaGlyHisPr 317
1418 AACGTCACGCTGACAGCAAGTTCGCACTCCCATCTGTAAGGCCACCC 1467
317 OAsnLeuCysValGlnValSerThrTTPGlyLysValGlnLeuGlnAlaC 334
1468 TAACTCTGTGTTCAGGTGAAACAGCTCGAGAACTCGAGCTGAGAGGT 1517
334 YSSerTTPAlaSPSerLeuGlyProPheLysAspAspMetLeuVal 350
1518 GCTTGTGGGCTGACTCTCTGGGCTCTCAAGACGATGTCTACTGTG 1567
351 GluMetLysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluProSe 367
1568 GAGACACGAGGCCCCCAGACAGATCCCTCTGTGCTTGGAAACCCAG 1617
367 TGLYcysThrProLeuProSerMetAlaSerThrArgAlaAlaArgLeuG 384
1618 TGGCTGTACTTACTTACCCAGCAAGCCCTCCAGC..... 1651
384 YGluGluLeuLeuGlnAspPheArgSerHisGlnCysMetGlnLeuTTP 400
1652 .....CTAAGG 1657
401 AsnAspAspAsnMetGlySerLeuTTPAlaCysProMetAspLysTyr11 417
1658 ...GACGATGACTTGGGAGCGCATATGGCCCTGCCCATGACAATAATCAT 1704
417 eHisArgArg 420
1705 CCACAAGCGC 1714
seq_name: gb_pat:AX073996
seq_documentation_block:

```

```

LOCUS AX073996 1753 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 4 from Patent WO0104304.
ACCESSION AX073996
VERSION AX073996.1 GI:12710228
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1753)
AUTHORS Presnell,S.R., Burkhead,S.K. and Pownder,S.L.
TITLE Human cytokine receptor
JOURNAL Patent: WO 0104304-A 4 18-JAN-2001;
Zymogenetics, Inc. (US)
FEATURES
source
location/Qualifiers
1..1753
/organism="Homo sapiens"
/db_xref="taxon:9606"
<2..1729
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28436.1"
/db_xref="GI:12710229"
/translacion="EPRNLSIQAVVLSFQVPTARCVLLEVOVPALVOFGS
VVYDCEPALGSFVRIWSTYQPRKEKLNHTQOUPALPMLVNSADGVHVLNSEE
OHEGLSLYMNVOGPPKPRMKNLTGPOITLNTHTDLYPCLOVMPLEPDSVRNIC
PEREDRAHONMOARILRLTLOSMLLDAPCSIPAPALCWRAPGCGCPVPLIS
MENWTVDVNSSEKLOECIMADSLGKLDDVLLERTGRPDNNSKLAEBSGCTSLP
SKASTRAARIGETLDDLOSQCQQLDDLDIGALUMACPMDXY IHRKRLVWLACILFA
AALSLILLLKDKAKGMLRLKQDVRSGAARAGRALILLSADSGFERLIGALASAL
COLPLVAVDLMRSRELISAGPVAMFHAQROAGEGVVVLLFSPGAVALCSERLOD
GVSGPAHGPHDAFRASLSCVLPDPELRAPGSYVGCDFRLTLPDAVPLFRTVPF
TLPSPDPLFGALQDPRAPRSGRHQERAEVSRALQPALDSYRHPPTPARGRVPG
AGRGADGT"
BASE COUNT 287 a 384 c 548 g 334 t
ORIGIN
alignment_scores:
Quality: 1154.50 Length: 340
Ratio: 4.260 Gaps: 3
Percent Similarity: 79.706 Percent Identity: 65.588
alignment_block:
US-09-899-471-2_COPY_21_440 x AX073996
Align seg 1/1 to: AX073996 from: 1 to: 1753
90 GlnGluSerArgAsnAlaSerLeuGlnAlaGlnValValLeuSerPheG1 106
2 GAGAGCGCTAGAGATGCTCTCTCCAGGCCCAAGTCGCTCTCTCTCCA 51
106 nAlaTyrProIleAlaArgCysAlaLeuLeuGlnValValGlnValProAla 123
52 GGCCTACCTACTGCCCCCTGCTCTCTGAGAGGTGAAAGTGGCTGTG 101
123 sPLeuValGlnProGlyGlnSerValGlySerAlaValPheAspCysPhe 139
102 CCGTGTGCAAGTTGGTGTGCTGTGGGCTCTGTGTATATACATGCTTC 151
140 GluAlaSerLeuGlyAlaGlnValGlnIleTTPSerTyrTTPLysProAr 156
152 GAGCGTCCCTAGGAGAGAGTACGATGTGCTCTTACTACACGCCAG 201
156 gTYrGlnLysGlnLeuAsnLeuThrGlnGlnLeuPro..... 168
202 GTACGAGAAAGAACTCAACACACAGCTGCTGCCCTGCCCTGGC 251
169 .....AspGlyAspAsnValLeuLeuThrLeuAspValSer 180
252 TCAACGTGTGACAGATGTGTGACAAAGTGCATCTGTCTGTAATGTCTT 301
181 GluGluGlnAspPheSerPheLeuLeuTyrLeuArgProValProAspAl 197

```



```

|||||  |||:|||||  |||  |||
302 GAGAGACAGACATTGGCCCTCTCCCTTACTGGAGATCAGGTCACAGGCC 351
197 aleuLysSerLeuTrpTyrIleAsnLeuThrGlyProGlnAsnIlePrl 214
352 CCCAAACCCCGGTGGACAAAACCTGACTGAGCCGACGATCATCTACT 401
214 euAsnHisThrAspLeuValProCysLeuCysIleGlnValTrpSerLeu 230
402 TGAACACACAGACCTGGTCCCTGCTGTATTCAGTGTGGCCCTGTG 451
231 GluProAspSerGluIuArgValGluPheCysProPheArgIuAspProG 247
452 GAACCTGACTCCGTTAGGACGAAACATCTGCCCTTCAGGAGAGGCCCG 501
247 yalaHisArgAsnLeuTrpHisIleAlaArgLeuArgValLeuSerPro 264
502 CGGACACACAGAACCTCTGGCAGCGCCGCCGACTGCTGACCCCTGC 551
264 lYValTrpGlnLeuAspAlaProCysCysLeuProGlyIleValThrLeu 280
552 AGAGCTGGCTGTGAGACGACCGCTGCTGCCGCCGAGAGAGGCGACTG 601
281 CysTrpGlnAlaProAspGlnSerProCysGlnProLeuValProProva 297
602 TGTGCGCGGCTCCGGTGGGACCCCTGCCAGCCACTGTGCCACCCCT 651
297 lProGlnLysAsnAlaThrValAsnGluProGlnAspPheGlnLeuVal 314
652 TTCCTGGGAGAACGTCCTGTGCAC..... 676
314 laGlyHisProAsnLeuCysValGlnValSerThrTrpGluLysValGln 330
677 .....GTGAACAGCTGGAGAGAGCTGCAG 700
331 leuGlnAlaCysSerTrpAlaAspSerLeuGlyProPheLysAspAspMe 347
701 CTGCAGAGAGCTGTGTGGCTGACTCCCTGGGCTCTCAAAAGAGATGT 750
347 lLeuValGlnMetLysThrGlyLeuAsnAspThrSerValCysAlaL 364
751 GCACCTGTTGGACACAGAGCCGCCGAGACACAGATCCCTGTGTGCT 800
364 euGluProSerGlyCysThrProLeuProSerMetAlaSerThrArgAla 380
801 TGAACCCAGCTGCTGACTTCACTACCAAGCCTCCACAGAGGCA 850
381 AlaArgLeuGlyGluLeuLeuGlnAspPheArgSerHisGlnCysMe 397
851 GCTCGCCCTGGAGAGTACTTACCAAGACCTGCAGTCCAGGCCAGTGTCT 900
397 tGlnLeuTrpAsnAspAspAsnMetGlySerLeuTrpAlaCysProMet 414
901 GCAGCTATGG...GACATGACTTGGAGCGCTATGGGCTGCCCCATGG 947
414 sPLysTyrIleHisArgArg 420
948 ACAATATACATCCACAGCGC 967

seq_name: gp_pat:AX073995
seq_documentation_block:
LOCUS AX073995 2076 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 3 from Patent WO0104304.
ACCESSION AX073995
VERSION AX073995.1 GI:12710227
KEYWORDS
SOURCE
ORGANISM synthetic construct.
synthetic construct
artificial sequence.
REFERENCE
AUTHORS 1 (bases 1 to 2076)
Presnell,S.R., Burkhead,S.K. and Powder,S.L.
TITLE Human cytokine receptor

```

```

JOURNAL Patent: WO 0104304-A 3 18-JAN-2001;
FEATURES
source Location/Qualifiers
1..2076
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="This degenerate sequence encodes the amino acid
sequence of SEQ ID NO:2."
misc_feature
1..2076
/note="n = A,T,C or G"
variation
1..2076
/note="N is any nucleotide"
BASE COUNT 232 a 273 c 399 g 265 t 907 others
ORIGIN

alignment_scores:
Quality: 1148.00 Length: 431
Ratio: 4.000 Gaps: 3
Percent Similarity: 66.589 Percent Identity: 50.116

alignment_block:
US-09-899-471-2_copy_21_440 x AX073995 ..

Align seq 1/1 to: AX073995 from: 1 to: 2076

2 GluArgLeuMetGluProGlnAspThrAlaArgCysSerLeuGlyLeu 18
111 ..... |||||||:|||||:|||||:|||||:|||||: 111
64 GARGNNTTNGTNGNCNCARAGATGCNACNCAYTITWMSCKGNGTNTWS 113
18 rCyHisLeuTrpAspGlyAspValLeuCysLeuProGlySerLeuGln 35
114 NTGYMNTVNTNGGAYWNSGAYATHYTMGYTNCNGNGNAGATHTGTC 163
35 eAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeuVal 51
164 CNGCNCNGCNCNGCNGTNTNGCNCNACNCAVYTNCARCNARGNTNGTN 213
52 LeuArgCysProGlnLysThrAspCysAlaLeuArgValArgValVal 68
214 YTMNGNTGYCARAARGARACAGATITGYCATYTNITGYTNMGNTNGCT 263
68 lHisLeuAlaValHisGlyHisTrpAlaGluProGlnGlu..... 81
264 NCAYTNTGNCNTNCAYGNCAYTGGARGARCNARGARAGARGARAAT 313
82 .AlaGlyLysSerAspSerGluLeuGlnGluSerArgAsnAlaSerLeu 97
314 TYGNGNGCNGCNGCNGAYWNSGNGTNGARGARCCNMGNAAAYGCNMSNTYN 363
98 GlnAlaGlnValAlaLeuSerPheGlnAlaTyrProIleAlaArgCysAl 114
364 CARCNCARGTNGTNTYTNWSNTTYCARCNCNTAYCCNACNGCNMGNTGYT 413
114 aleuLeuGlnValGlnValProAlaAspLeuValGlnTrpGlyGlnSerV 131
414 NYTNTNGARGTNCARGTNCNGCNGCNYTNTGTCARTTYGCMARMSNG 463
131 aGlySerAlaValPheAspCysPheGlnAlaSerDeuGlyAlaGluVal 147
464 TNGCNSNGTNGTNTAYGAYTYTYGARGCNGCNYTNGGNSNGARGTN 513
148 GlnIleTrpSerTyrThrLysProArgTyrGlnLysGluLeuAsnLeuTh 164
514 MGNAYHTGGMSNTAYACNARCCNMGNMTAYGARRARARGAYTNAACAYAC 563
164 rGlnGlnLeuPro.....AspGlyAspA 172
564 NCARCARYTNCNCNTNTCCNTGGYTNAAGTNTMSNCCNAGYGGNGATA 613
172 snValLeuLeuThrLeuAspValSerGluGluGlnAspPheSerPheLeu 188
614 AYTNCAYTYTNGTNTYTNAAAYGTNTMSNGARGARCARCAYTYGNTYTN 663

```


189 LeuTYrLeuArGProValProAspAlaLeuYsSerLeuTYrLYsAs 205
 664 YrTNAAYGAAYCARGTNCARGNCCNNAARCCNMGWGTGCATAARA 713
 205 nLeuThGlyProGlnAsnIleThrLeuAsnHisThrAspLeuValProC 222
 714 YrTNAAGNCCNCArATHACNTNAAYCAVACNGAYYNGTNCNT 763
 222 YsLeuCYsIleGlnValIleThrSerLeuGluProAspSerGluArgValGlu 238
 764 GYrTNTGArTHCARGTNTGCCNTNGARCCGAYWSGTNGNACNAAY 813
 239 PheCYsProPheArgGluAspProGlyAlaHisArgAsnLeuTYrHisI 255
 814 ATHTGTCNTTYMGARGAYCCMNGCCNCAVCARAAYTNTGGCARGC 863
 255 eAlaArgLeuArgValLeuSerProGlyValIleProGlnLeuAspAlaProC 272
 864 NGCNMGNTNMGNNTYNTNACNTNCArWSNTGTNTNGAYGCNCCNT 913
 272 YsCYsLeuProGlyValIleThrLeuCYsTrpGlnAlaProAspGlnSer 288
 914 GYrSNTNCCNCCNGARCCNCCNTTGTGTGCMGNCNCCNCGNGNGAY 963
 289 ProCYsGlnProLeuValProProValProGlnYsAsnAlaThrValAs 305
 964 CCNTGTCARCCNNTNGTNCNCCNTTMSNTGGARAAYGTNACNCTNA 1013
 305 nGluProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuCYsValG 322
 1014 YARGTNTGTGARTTYCCNTNTYNTNAARGNCAVCCNNAAYTNTGTGTNC 1063
 322 InValSerThrTrpGluValGlnLeuGlnAlaCYsSerTrpAlaAsp 338
 1064 ARGTNAAYWSNMSGARARARTNCARTNCARGARTGYTNTGGCCNGAY 1113
 339 SerLeuGlyProPheLYsAspAspMetLeuValGluMetLYsThrG 355
 1114 WSNYTNCGNCCNTNAARGAYGAYGTNTNTNTNGARACNMGNGNCC 1163
 355 YLeuAsnAspThrSerValCYsAlaLeuGluProSerGlyCYsThrProL 372
 1164 NCARGAYAAATGMWSNTNTGTGNTNGARCCNWSNCGNTGTACNWSNY 1213
 372 euProSerMetAlaSerThrArgAlaAlaArgLeuGluGluLeuLeu 388
 1214 TNCNMSNARCCNWSNACNMGNCNCGNMGNTNGNGARATYTTNTYN 1263
 389 GlnAspPheArgSerHisGlnCYsMetGlnLeuTrpAsnAspAspAsnMe 405
 1264 CARGAYTNCARFNSNGNCARTGYTNCARTNTGG...GAYGAGAYTT 1310
 405 tGlySerLeuTrpAlaCYsProMetAlaPLYsTYrIleHisArg 419
 1311 NGNNGCNTYNTGTGCCNTGCCNATGAGAYARTATATCAVAA 1353
 seq_name: gp_pat:AX350975
 seq_documentation block:
 LOCUS AX350975 2109 bp DNA linear PAT 06-FEB-2002
 DEFINITION Sequence 9 from Patent WO0190358.
 ACCESSION AX350975
 VERSION AX350975.1 GI:18616351
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified
 unclassified.
 REFERENCE 1 (sites)
 AUTHORS Gorman,D.M.
 TITLE Mammalian receptor proteins: related reagents and methods
 JOURNAL Patent: WO 0190358-A 9 29-NOV-2001;
 SCHERING CORPORATION (US)

FEATURES
 source Location/Qualifiers
 1..2109
 /organism="unidentified"
 /db_xref="taxon:32644"
 /note="primate: surmised Homo sapiens"
 BASE COUNT 234 a 267 c 415 g 269 t 924 others
 ORIGIN
 alignment_scores:
 Quality: 1076.00 Length: 446
 Ratio: 3.884 Gaps: 4
 Percent Similarity: 62.108 Percent Identity: 46.637
 alignment block:
 US-09-899-471-2_COPY_21_440 x AX350975 ..
 Align seg 1/1 to: AX350975 from: 1 to: 2109
 2 GluTrpGluMetGluProGlnAspThrAlaArgCYsSerLeuLeu 18
 64 GARGNTNTNGTNGCCNCCNCArGAYGCNCAVGTGWSNCCNCGNTTWS 113
 18 rCYsHisLeuTrpAspGlyAspValLeuCYsLeuProGlySerLeuGln 35
 114 NTGYGNTNTGGGAYWSGAVATHTNTGYTNTCCNCGNGATATGCTNC 163
 35 eAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeuVal 51
 164 CNGCNCCNCGNCCNCTNTNGCNCCNACNCAVYTNCArACNARGARTNTG 213
 52 LeuArgCYsProGlnLYsThrAspCYsAlaLeuArgValArgValAla 68
 214 YTNNGNTYCARAARGARCCNCAVGTGYTNTGYTNTMGNTNGCNT 263
 68 HisLeuAlaValHisGlyHisTrpAlaGluProGluGlu..... 81
 264 NCAYTNTCCNCTNCAVGCNCAVGTGGARARCCNARGAYGARGARART 313
 82 ..AlaGlyLYsSerAspSerGluLeuGlnGluSerArgAsnAlaSerLeu 97
 314 TYGNGNCGNCCNCAVYTNNGTNGTNGARCCNMGNAAYGCNWSNTYN 363
 98 GlnAlaGlnValValLeuSerPheGlnAlaIleProIleAlaArgCYsAl 114
 364 CARGCNCArGNTNGTNTWSNTTYCARCGCNTAYCCACNCCNNGNTGYGT 413
 114 AlaLeuGlnValAlaGlnValProAlaAspLeuValGlnProGlyGlnSer 131
 414 NYTNTNGARGTNCARGTNCNCGNCCNCGNCTNTNCARTTYGNCARWSNG 463
 131 aGlySerAlaValPheAspCYsPheGluAlaSerLeuGlyAlaGluVal 147
 464 TNGCNWSNGTNTAYGAYGTYTGYGARGCNCCNCTNGNWSNGARGTNT 513
 148 GlnIleTrpSerTYrThLYsProArgTYrGlnLYsGlnLeuAsnLeuTh 164
 514 MGNATHGMSNTATACNCArCCNMGNTATYGAR...ARTTNAAYCAVAC 563
 164 rGlnGlnLeuPro..... 168
 564 NCARCArTNCNCAVGTGYMGNGNTNGARGTNTGGAAYWSATHCCNN 613
 169AspGlyAspAsnVal 173
 614 SNTGYTGGCNTNCCNTGGYTNAAYGTNWSNCGNAGAYGNGAYAGTNT 663
 174 LeuLeuThrLeuAspValSerGluGluGlnAspPheSerPheLeuLeu 190
 664 CAYTNTGNTNTNAAGTWNMSGARCARCAVYTYGNTTMSNTNTTA 713
 190 rLeuArgProValProAspAlaLeuLYsSerLeuTrpTYrLYsAsnLeu 207
 1 1


```

714 YTGGAATCAGTCAGCGCCGCCNCCNARCCNMGTCGGCAVAAATYTTA 763
207 hrglyProGlnAsnIleThrLeuAsnHisThrAspLeuValProCysLeu 223
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
764 CNGSCGNCARATHTATACATYTTAAACACACACATYTTGTCGTCGTTN 813
224 CysIleGlnValTyrSerLeuGluProAspSerGluArgValGluPheCys 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
814 TGYATHCARGTGTGGCCNTGARCNCAGATWSNGTMMGNACNAATHTG 863
240 sProPheArgGluAspProGlyAlaHisArgAsnLeuTyrHisIleAla 257
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
864 YCCNTTITMGNGARGAYCCNMGNCNCAYCAARAAYTTTGCCAGCNCNC 913
257 rGleuArgValLeuSerProGlyValTyrGlnLeuAspAlaPheCysCys 273
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
914 GNTTNGMTTNTTACNTTNCARWSNTGTTTNTGATGCGCCNTGWSN 963
274 LeuProGlyValThrLeuGlyTyrGlnAlaProAspGlnSerProCys 290
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
964 YTNCCGNCNGARGCNCNTTNTGTYGMCNGCNCNGCNCNGCNCATCCNTG 1013
290 sGlnProLeuValProProValProGlnIlyAsnAlaThrValAsnGlu 307
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1014 YCARCCNTTNTGCCNCCNTTMSMTGCGARAAYTTNACNGTNGAY.... 1059
307 rGlnAspPheGlnLeuValAlaGlyHisProAsnLeuGlyValGlnVal 323
1060 .....GTN 1062
324 SerThrTyrGlnIlyValGlnLeuGlnAlaCysSerTyrAlaAspSerle 340
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1063 AAATMSMNGARAATYTCARTNCARCARGTCTYTTGGCCNCAATMSNTY 1112
340 uGlyProPheIlyAspAspMetLeuLeuValGluMetIlyThrGlyLeu 357
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1113 NGGCCNTTNAARAGAYGTGYTNTTNTNGARCCNMGNCNCNCARG 1162
357 sNAsnThrSerValCysAlaLeuGluProSerGlyCysThrProLeuPro 373
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1163 AATATMGMSNTTNTGTCNTTNGARCCNMSNGNTGYACNMSNTYTCN 1212
374 SerMetIleAspThrArgAlaAlaArgLeuGlyGluGluLeuGlnAsn 390
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1213 MSNARCCNMSNACNMGNCNCNNGNTNGNGATATYTTNTNCARCA 1262
390 pPheArgSerHisGlnCysMetGlnLeuTyrAsnAspAspAspMetGlyS 407
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1263 YTTNCAKMSNCGNCARTGYTTCARATYTTG...GAYGAYGAYTTNGNC 1309
407 eLeuTyrAlaCysProMetAspIlyTyrIleHisArg 419
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1310 CNTTNTGGCCNTGCCNATGGATATATATACATTAAR 1347

seq_name: gb_pat:AX073998

seq documentation block:
LOCUS AX073998 1725 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 6 from Patent WO0104304.
ACCESSION AX073998
VERSION AX073998.1 GI:12710230
KEYWORDS
SOURCE
  ORGANISM
    synthetic construct.
    artificial sequence.
  REFERENCE
    1 (bases 1 to 1725)
    Presnell,S.R., Burkhead,S.K. and Powder,S.L.
    Human cytokine receptor
    Patent: WO 0104304-A 6 18-JAN-2001;
    JOURNAL
      ZymoGenetics, Inc. (US)
      Location/Qualifiers
        1..1725
        /organism="synthetic construct"

```

```

/db_xref="taxon:32630"
/note="This degenerate sequence encodes the amino acid
sequence of SEQ ID NO:5."
misc-feature
  1..1725
  /note="n = A,T,C or G"
variation
  1..1725
  /note="N is any nucleotide."
BASE COUNT 192 a 226 c 332 g 216 t 759 others
ORIGIN

```

alignment_scores:

Quality:	Ratio:	Length:
852.50	3.947	339
Percent Similarity:	Gaps:	Percent Identity:
63.717	3	48.968

alignment_block:

US-09-899-471-2_COPY_21_440 x AX073998

Align seg 1/1 to: AX073998 from: 1 to: 1725

```

90 gInGluSerArgAsnAlaSerLeuGlnAlaGlnValValLeuSerPheG1 106
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 GARGARCCNMGNAAYCGNMSNTTNCARCCNARCCNARCCNARCCNARCCN 50
106 nAlaTyrProIleAlaArgCysAlaLeuLeuGluValGlnValProAla 123
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 RGCNTAAYCCNACNCGNMGNTGYGTNTTNTNGARCCNARCCNARCCN 100
123 sPLeuValGlnProGlyGlnSerValGlySerAlaValPheAspCysPhe 139
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 CNTTNTNCARATYTTGGNCARMSNGTNGNMSNGTNGTNTAATGATGYT 150
140 GluAlaSerLeuGlnIlyAlaGlnIleTyrSerTyrThrIlyProAr 156
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151 GARGCNCGYTTNGNMSNGARGTGNTGMSNTATGMSNTATACNARCCNMG 200
156 gTyrGlnIlyGluLeuAsnLeuThrGlnGlnLeuPro..... 168
201 NTAYGARAARGAYTTNAACAYACNARCCARATYTCNCGNTTNCNTG 250
169 .....AspGlyAspAsnValLeuLeuThrLeuAspValSer 180
251 TNAAYGTNMSNCGNAYGGNGAYAAATGTCATYTTGATYTTNAATGWSN 300
181 GluGlnGlnAspPheSerPheLeuLeuTyrLeuArgProValProAspAl 197
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 GARGARCARCATYTTGCGNTTMSNTTNTATGGAAYCARCINCARCNC 350
197 aLeuIlySerLeuTyrTyrIlyAsnLeuThrGlyProGlnAsnIleThr 214
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
351 MCCNARCCNMGNTGGCAVAAARAAYTTNACNCGNCNCARATHTACNY 400
214 eUAsnHisThrAspLeuValProCysLeuCysIleGlnValTyrSerLeu 230
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
401 TNAATCAACACATYTTGTCNTGTTTGTATACAGTNTGCCNTYTN 450
231 GluProAspSerGluArgValGluPheCysProPheArgGluAspProG1 247
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
451 GARGCNGAYWSNGTMMGNACNAATHTGCTTTTNGNARARAYCCNMG 500
247 yAlaHisArgAsnLeuTyrHisIleAlaArgLeuArgValLeuSerProG 264
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
501 NGCNCAYCARAAYTTNTGGCARGCNCNNGNTTNGMTNTYTTNACNTNC 550
264 IlyValTyrGlnLeuAspAlaProCysCysLeuProGlyIlyValThrLeu 280
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
551 ARMSWTGTYTNTNGATGNCNTGTGWSYTTTNCNCGNARGCNCNTYTN 600
281 CysTyrGlnAlaProAspGlnSerProCysGlnProLeuValProProVa 297
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
601 TGYTGMGNCNCNCGNCGNCGNAYCCNTGYCARCCNAYTTNGTNCNCNTY 650

```


297 lProGlnLysAsnAlaThrValAsnGluProGlnAspPheGlnLeuValA 314
 651 NMSNTGGGARAAYGTNACNGTNGAY 675
 314 lAglyHisProAsnLeuGlyValGlnValSerThrTyrGlyValGln 330
 676 699
 331 LeuGlnAlaCysSerThrPalaAspSerLeuGlyProPheLysAspAsp 347
 700 YTNCAAGARFGYTYNTGGCGAGATVMSNTNGNCCNTYTAAGATATGCT 749
 347 tLeuLeuValGlnMetLysThrGlyLeuAsnAsnThrSerValCysAla 364
 750 NYTNYTYTNARACNMNGNCCNCAAGATVAAVGMNMSNTNTGYGCAN 799
 364 euGluProSerGlyCysThrProLeuProSerMetAlaSerThrArgAla 380
 800 TNARCCNMSNGNTGTACNMSNTTCCNMSNAAGCNCNMSNACNMNGCN 849
 381 AlaArgLeuGlyGluGlnLeuLeuGlnAspPheArgSerHisGlnCys 397
 850 GCNMGNTNGNGARATAYTNTNCAAGATVTAACNMSNGNCAAGCTT 899
 397 tGlnLeuTyrAsnAspAspAspMetGlySerLeuTyrPalaCysProMet 414
 900 NCARNTYTG...GAYGAYGAYTNGNCCNTNTNGGCGTGCYCCNATG 946
 414 sPlyTyrIleHisArg 419
 947 AYARATAYATHCAATAR 963
 seq_name: gb_pat:AX136589

seq_documentation_block:
 LOCUS AX136589 694 bp DNA linear PAT 30-MAY-2001
 DEFINITION Sequence 511 from Patent EP1067182.
 ACCESSION AX136589
 VERSION AX136589.1 GI:14272993
 KEYWORDS

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 694)
 AUTHORS Ota,T., Isegar,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and Hayashi,K.
 TITLE Secretory protein or membrane protein
 JOURNAL Patent: EP 1067182-A 511 10-JAN-2001;
 Helix Research Institute (JP)
 FEATURES
 source 1..694
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 115 a 214 c 219 g 141 t 5 others
 ORIGIN

Alignment_scores:
 Quality: 464.00 Length: 142
 Ratio: 3.803 Gaps: 1
 Percent Similarity: 85.915 Percent Identity: 69.014

Alignment block:
 US-09-899-471-2_COPY_21_440 x AX136589

Align seg 1/1 to: AX136589 from: 1 to: 694
 1 LeuGlnArgLeuMetGluProGlnAspPheAlaArgCysSerLeuGlyLe 17
 268 CTGAGAGAGGCTTGTGGGCGCTCAGAGCGCTACCCACTCTCTCCGGGCT 317
 17 uSerCysHisLeuTyrAspGlyAspValLeuGlyLeuProGlySerLeuG 34

|||||.....
 318 CTCCTGGCGCCCTCTGTGGACAGTACACTCTGCTCCCTGGGACACTGC 367
 34 lnsrAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeu 50
 368 TGCCCTGCTCGGCGCCCTGTGCTGGCCCTACGACCTGCAGACAGACTG 417
 51 ValLeuArgCysProGlnLysThrAspCysAlaLeuArgValArgVala 67
 418 GTGCTGAGTGCACAGAGAGACCGACTGTGACTCTGTCTGCTGTGGC 467
 67 lValHisLeuAlaValHisGlyHisTyrPalaGluProGluGlu..... 81
 468 TCTCCACTGTGGCGCTGCATGGCAGCTGGAGAGCCCTGAAGATAGGAAA 517
 82AlaGlyLysSerAspSerGluLeuGlnGlnSerAlaArgAsnAlaSer 96
 518 ACTTGGAGAGGACGCTGACTCANGGTGGAGAGAGCTTANGAATGCTCT 567
 97 LeuGlnAlaGlnValValLeuSerPheGlnAlaTyrProIleAlaArgC 113
 568 CTCACAGGCCCAAGTCTGCTCTCTCTACAGCGCTACCTACTGCCCCGCTG 616
 113 sAlaLeuLeuGlnValGlnValProAlaAspLeuValGlnProGlyLns 130
 617 CCTCTGCTGTGANGTCAAGTCCCTGCTGCCTGCTGCTGCTGCTGCTGCTG 665
 130 eValGlySerAlaValPheAspCys 138
 666 CTGTGGGCTCTGTGTATATGACTGC 691
 seq_name: gb_htg:AC026196

seq_documentation_block:
 LOCUS AC026196 191834 bp DNA linear HTG 29-MAY-2000
 DEFINITION Homo sapiens chromosome 3 clone RP11-474F16 map 3p, WORKING DRAFT
 SEQUENCE, 22 unordered pieces.
 ACCESSION AC026196
 VERSION AC026196.3 GI:8101230
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 191834)
 AUTHORS Cai,T., Dong,X., Gao,Q., Gu,J., Gong,J., He,F., Kang,N., Lu,T., Ma,Q., Rong,L., Shen,Y., Tan,X., Wang,H., Xi,Y., Xu,Y., Yao,Z., Zheng,Z., Zhu,N., Zhou,X., Zhou,Y., Qiang,B., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H., Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Liu,W., Li,W., Li,Y., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,Y., Tao,R., Wang,H., Wang,Y., Wang,J., Wang,H., Wang,L., Wang,R., Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Yu,J. and Yang,H.
 TITLE Chromosome 3p genomic sequence
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 191834)
 AUTHORS Wang,R., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B., Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y., Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y., Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D., Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J. and Yang,H.

TITLE Direct Submission
 JOURNAL Submitted (21-MAR-2000) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China
 COMMENT On May 29, 2000 this sequence version replaced gl:8050894.
 -----Genome Center

Center: Beijing Center
Center code: Beijing
Website: http://hgci.igtp.ac.cn
http://www.genomics.org.cn
Contact: hgci@igtp.ac.cn

Project Information
Center project name: 13 project
Center clone name: RP11-474F16

Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; ET 5% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 185328 bases at least Q40
Consensus quality: 192037 bases at least Q30
Consensus quality: 196045 bases at least Q20
Insert size: 179116; sum-of-contigs
Quality coverage: 5.75x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

```
1 1632: contig of 1632 bp in length
* 1633 1732: gap of unknown length
* 1733 2953: contig of 1221 bp in length
* 2954 3053: gap of unknown length
* 3054 4934: contig of 1881 bp in length
* 4935 5034: gap of unknown length
* 5035 7618: contig of 2584 bp in length
* 7619 7719 9805: contig of 2087 bp in length
* 9806 9905: gap of unknown length
* 9906 12859: contig of 2934 bp in length
* 12860 12959: gap of unknown length
* 12960 16605: contig of 3646 bp in length
* 16606 16705: gap of unknown length
* 16706 19419: contig of 2714 bp in length
* 19420 19519: gap of unknown length
* 19520 22470: contig of 2951 bp in length
* 22471 22570: gap of unknown length
* 22571 27404: contig of 4834 bp in length
* 27405 33396: contig of 5892 bp in length
* 33397 33496: gap of unknown length
* 33497 39063: contig of 5567 bp in length
* 39064 39163: gap of unknown length
* 39164 47186: contig of 8023 bp in length
* 47187 47286: gap of unknown length
* 47287 54806: contig of 7520 bp in length
* 54807 54906: gap of unknown length
* 54907 62077 62176: contig of 7170 bp in length
* 62077 62177 73474: contig of 11298 bp in length
* 73475 73574: gap of unknown length
* 73575 83952: contig of 10378 bp in length
* 83953 84052: gap of unknown length
* 84053 99282: contig of 15230 bp in length
* 99283 99382: gap of unknown length
* 99383 114261: contig of 14879 bp in length
* 114262 114361: gap of unknown length
* 114362 129674: contig of 15313 bp in length
* 129675 129774: gap of unknown length
* 129775 144971: contig of 15197 bp in length
* 144972 145071: gap of unknown length
* 145072 191834: contig of 46763 bp in length.
```

Location/Qualifiers
1. 191834
/organism="Homo sapiens"

```
misc_feature /db_xref="taxon:9606"
misc_feature /chromosome="3"
misc_feature /map="3p"
misc_feature /clone="RP11-474F16"
1. 1632
misc_feature /note="assembly_name:Contig17"
1733. 2953
misc_feature /note="assembly_name:Contig18"
3054. 4934
misc_feature /note="assembly_name:Contig20"
5035. 7618
misc_feature /note="assembly_name:Contig22"
7719. 9805
misc_feature /note="assembly_name:Contig23"
9906. 12859
misc_feature /note="assembly_name:Contig24"
12960. 16605
misc_feature /note="assembly_name:Contig25"
16706. 19419
misc_feature /note="assembly_name:Contig26"
19520. 22470
misc_feature /note="assembly_name:Contig27"
22571. 27404
misc_feature /note="assembly_name:Contig28"
27505. 33396
misc_feature /note="assembly_name:Contig29"
33497. 39063
misc_feature /note="assembly_name:Contig30"
39164. 47186
misc_feature /note="assembly_name:Contig31"
47287. 54806
misc_feature /note="assembly_name:Contig32"
54907. 62076
misc_feature /note="assembly_name:Contig33"
62177. 73474
misc_feature /note="assembly_name:Contig34"
73575. 83952
misc_feature /note="assembly_name:Contig35"
84053. 99282
misc_feature /note="assembly_name:Contig36"
99383. 114261
misc_feature /note="assembly_name:Contig37"
114362. 129674
misc_feature /note="assembly_name:Contig38"
129775. 144971
misc_feature /note="assembly_name:Contig39"
145072. 191834
/note="assembly_name:Contig40"
BASE COUNT 48049 a 46926 c 46746 g 47961 t 2152 others
ORIGIN
alignment_scores:
Quality: 316.00 Length: 307
Ratio: 2.612 Gaps: 8
Percent Similarity: 39.414 Percent Identity: 31.270
alignment_block:
US-09-899-471-2_Copy_21_440 x AC026196/rev
Align seg 1/1 to reverse of: AC026196 from: 1. 191834
23 Aspglyaspvalleucysileuproclyserleuglnserlaspriolypf 39
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
38840 GACAGTACATACCTGCTGCTGGGAGACATCGTCCGCTCCGGGCC 38791
39 ovalleuvalprothraragleuglnfrhicleuvalleuargcysprog 56
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
38790 CGTCTGGCGCTACGACCTCCACAGACCTGAGCTGAGAGTCCACA 38741
56 InlysthraspcysalaleuargvalargvalvalahisileuAlaval 72
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
38740 AGGAGACCGACTGTGACCTCTGTCTGCGTGTGCTGCTCCACTTGCGG 38691
```


Align seg 1/1 to: AAS18132 from: 1 to: 2314

```

1 LeuGIuArGleuMetGIuProGIuAspThrAlaArgCysSerLeuGlyLe 17
259 CTGAGAGACGATGAGACCTCAGACACGCTGCTCTCTAGAGCCT 308
17 USeRCysHisLeuTrpAspGIuAspValLeuCysLeuProGIuSerLeu 34
309 CTCTGCACACTGAGGATGGTGAAGTGTGCTGCTGCTGAGAGCCTCC 358
34 InSeRAlaProGIuProValLeuValProThrArgLeuGIuThrGIuLeu 50
359 AGTCGTGCCCCGAGCCCTGTGCTAGTCCCTACCCGCTGCAGAGGAGCTG 408
51 ValLeuArGcYsProGIuLysThrAspCysAlaLeuArgValArgVala 67
409 GTGCTGAGGTTCACAGACAGACAGATGCGCCCTGTGCTGCTGTGGT 458
67 ValHisLeuAlaValHisGIuHisTrpAlaGIuProGIuGIuAlaGIu 84
459 GGTCACACTGGCCGTCATGGGCACTGGGCAAGCTGAAAGAGCTGGAA 508
84 YSeRAspSerGIuLeuGIuInSeRArgAsnAlaSerLeuGIuAlaGIu 100
509 AGTCGTATTCAGAACTCCAGAGAGTCTAGGAAGCCTCTCTCCAGGCCCA 558
101 ValValLeuSerPheGIuAlaTrpProIleAlaArgCysAlaLeuLeu 117
559 GTGGTGTCTCTCTCCAGGCTACCCCAATGCCCTGTGCTGCTGTGCTGA 608
117 uValGIuValProAlaAspLeuValGIuProGIuGIuInSeRValGIuSe 134
609 GGTCACAGTCCCGCTGACCTGCTGACCTGCTGCTGCTGCTGCTGCTG 658
134 lAvalPheAspCysPheGIuAlaSerLeuGIuAlaGIuValGIuInLe 150
659 CGGTATTTGACTGTTTCAGGCTAGTCTTGGGGCTGAGGTACAGATCTG 708
151 SerTrpThrLysProArgTrpGIuGIuLysGIuLeuAsnLeuThrGIuIn 167
709 TCCTACACGAGACCCAGGTACAGAAAGACTCAACCTCACACAGACGCT 758
167 uPro..... 168
759 GCCTGACTGACGAGGGCTCTGAATCCGGAGACAGCTCCAGAGCTGCTG 808
169 .....AspGIuAspAsnValLeuLeuThr 176
809 TCCTGCCCTGGCTCAATGTGTCTACAGATGGTGACAAATGTCCTTGACA 858
177 LeuAspValSerGIuGIuGIuAspPheSerPheLeuLeuTrpLeuArgPr 193
859 CTGGATGTCTCTGAGGAGCAGGACTTACCTTCTTACTGACTCCGCTCC 908
193 oValProAspAlaLeuLysSerLeuTrpTrpLysAsnLeuThrGIuPro 210
909 AGTCCGCGATGCTCTCAATCTCTGTGTACAAAACCTGACTGAGACTTC 958
210 lAnsnIleThrLeuAsnHisThrAspLeuValProCysLeuCysIleGIu 226
959 AGAACATTAATTAAACACACACAGACTGTGCTCCCTGCTGCTGACTTAC 1008
227 ValTrpSerLeuGIuProAspSerGIuArgValGIuLysPheCysProPhe 243
1009 GTGTGGTGGTACAGGCTGAGAGGCTGCAATTTCTCCCTTCCTCCG 1058
243 gGIuAspProGIuAlaHisArgAsnLeuTrpHisIleAlaArgLeuArgV 260
1059 GGAAGATCCCGGTGACACAGAACCTCTGCAATACAGGCTGCGCGG 1108
260 AlLeuSerProGIuValTrpGIuLeuAspAlaProCysCysLeuProGIu 276
1109 TACTGTCCCGAGGGGTATGCGAGCTAGATGCGCTGCTGCTGCGGGGC 1158

```

```

277 LysValThrLeuCysTrpGIuAlaProAspGIuSerProCysGIuProLe 293
1159 AAGGTAAACACTGTGTGGCAGGACACAGACAGATCCCTGCGAGCCACT 1208
293 uValProProValProGIuLysAsnAlaThrValAsnGIuProGIuAsp 310
1209 TTGTCCACCACTGCCCCAGAGAACGCTGCTGTGAATGACCCACAAAT 1258
310 heGIuLeuValAlaGIuHisProAsnLeuCysValGIuAlaSerThrTrp 326
1259 TCCAGTGTGTGGAGGACCAOCCACACTGTGTCCAGGTGAGACCTGG 1308
327 GIuLysValGIuLeuGIuAlaCysSerTrpAlaAspSerLeuGIuProPh 343
1309 GGAAGGTTCACCTGCACAGCTGTGTGTGGCTGACTCTTGGGGCCCTT 1358
343 eLysAspAspMetLeuLeuValGIuMetLysTrpGIuLeuAsnThrS 360
1359 CAAGATGATATGCTGTGTGTGAGATGAATAACGGCTCAACAAACAT 1408
360 eValCysAlaLeuGIuProSerGIuCysThrProLeuProSerMeta 376
1409 CAGTGTGTGCTTGGAAACCAAGTGGCTGTACACCACTGCCAGCATGGCC 1458
377 SerThrArgAlaAlaArgLeuGIuGIuGIuLeuLeuGIuAspPheArgSe 393
1459 TCCACGAGAGCTGTGCTGCTGGAGAGAGATGCTGCACAACTCCGATC 1508
393 rHisGIuCysMetGIuLeuTrpAsnAspAspAsnMetGIuSerLeuTrpA 410
1509 ACACAGTGTATGACAGCTGTGGAACGATGACAAATGAGATGCTATGGG 1558
410 lAcYsProMetAspLysTrpIleHisArgArg 420
1559 CTGCGCCATGAGACAAATCATCCACAGGCGC 1590
seq_name: /SID1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: AAS18133
seq_documentation_block:
ID AAS18133 standard; cDNA; 2094 BP.
AC AAS18133;
AT
CT 26-MAR-2002 (first entry)
DE Mouse DCRS7 reverse translation generic cDNA.
KW Mouse: DNAX cytokine receptor subunit 7; DCRS7; phosphate labelling; ss;
KW gene therapy; protein therapy; immunological disorder.
OS Mus musculus.
PN MO200190358-A2.
PD 29-NOV-2001.
PE 23-MAY-2001; 2001MO-US16767.
PR 24-MAY-2000; 2000US-206862P.
SCHE (SCHE) SCHERING CORP.
Gorman DM;
WPI; 2002-106198/14.
Isolated antigenic human or mouse DNAX receptor subunit-like
polypeptide useful for detecting antibodies generated in response to
presence of increased protein levels or immunological disorders -
Disclosure; Page 20-21; 148pp; English.

```


CC The invention relates to primate and rodent DNA cytokine receptor
 CC subunit (DCRS) polypeptides and the polynucleotides encoding them. The
 CC receptor, or their portions may be useful as phosphate labelling enzymes
 CC to label general or specific substrates. The subunits may also be
 CC functional immunogens to elicit recognising antibodies, or antigens
 CC capable of binding antibodies. A combination, e.g., including a DCRS can
 CC be used as an immunogen for the production of antisera or antibodies
 CC capable of distinguishing between other cytokine receptor family members.
 CC A purified DCRS can also be used as a reagent to detect antibodies
 CC generated in response to the presence of elevated levels of expression,
 CC or immunological disorders which lead to antibody production to the
 CC endogenous receptor. This sequence represents mouse DCRS7 reverse
 CC translation generic cDNA.

Sequence 2094 BP: 247 A; 251 C; 393 G; 285 T; 918 other;

Alignment scores:
 Quality: 1729.00 Length: 441
 Ratio: 5.056 Gaps: 1
 Percent Similarity: 77.551 Percent Identity: 70.975

Alignment block:

US-09-899-471-2-copy_21_440 x AAS18133 ..

Align seg 1/1 to: AAS18133 from: 1 to: 2094

```

2  GUAATGLeuMetGluProGlnAspThrAlaArgCysSerLeuGluLeu 18
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
64  GARGNTNATGARGCARGATACGACGCGMGNTGMSNTNGNTNWS 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
18  TCSHSLLeuTPAspGlyAspValLeuCysLeuProGlySerLeuGln 35
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
114  NTGCAYTNTGGAGYGNGAYGTNTNTGTYTNCNGGMSNTTNCARW 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
35  eRAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeuVal 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
164  SNGCCGCGNGCGTNTGNTGCCNACNMGNTTNCARACGARGTYTNGTN 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
52  LeuArgCysProGlnLysThrAspCysAlaLeuArgValArgValVal 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
214  YTNMGNTGCCCARARACARGATGTGCTNTGCTNGGNTNGT 263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68  LHSLeuAlaValHisGlyHisThrAlaGluProGluGluAlaGlyLys 85
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
264  NCAYTTCGNGTNCAYGNGCAYTGGCGNARGCNGARGCNGNARW 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
85  eRAspSerGluLeuGlnGluSerArgAsnAlaSerLeuGlnAlaGlnVal 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
314  SNGAYTWSGARGTNCARGARMSNMAAYGCMWSNNTNCARGCNCARGTN 363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
102  ValLeuSerPheGlnAlaTyrProIleAlaArgCysAlaLeuLeuGluVal 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
364  CTNTNMSNTTYTCARGCNTATCCNATGCGMGNTGCTTNTTNGARGT 413
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
118  IGLValProAlaAspLeuValGlnProGlyGlnSerValGlySerAlaVal 135
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
414  NCARGTCGCGNGCAYTNTGTCARCCNGCARGMSNGTNGGMSNCG 463
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
135  alphaPheCysPheGlnAlaSerLeuGlyAlaGluValGlnIleTyrSer 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
464  TMTTGGATGTYTTCARGCMTWSNNTNGCGNCARGTNCARATHTTGWSN 513
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
152  TyrThrLysProArgTyrGlnLysGluLeuAsnLeuThrGlnGlnLeuP 168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
514  TATACNAARCCMGNTATCARARARATYTAATYTAACNARCARATYTNCC 563
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
168  O..... 168
564  NGAYTGYMGNGNTNGARGTNMGNGAYWSNATHCARMSNTGTYGGCTNY 613
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
169  .....AspGlyAspAsnValLeuLeuThrLeu 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

614  TNCNTGCTNAAYGTNWSNACNGAYGNGAYAGTYNTNTNACNTYN 663
178  AspValSerGluGlnGlnAspPheSerPheLeuLeuTyrLeuArgProVal 194
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
664  GAGTWSNNGARGCARCARGATYTTWSNTTYTNTNTATYTTNMGCCNGT 713
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
194  IProAspAlaLeuLysSerLeuTyrTyrLysAsnLeuThrGlyProGlnA 211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
714  NCNGAYGCTYTAARMSNTNTGCTATTAARATYTAATNCCNGCCNAR 763
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
211  snLThrLeuAsnHisThrAspLeuValProCysLeuCysIleGlnVal 227
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
764  AYTHACNTNAAVCAYACNGAYTTCNCCNTGTYTNTGATATHCARGTN 813
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
228  TrpSerLeuGluProAspSerGluArgValGluLeuPheCysProPheArg 244
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
814  TGGMSNTNGARCCNGAYWSNGARGNGTNGARTTYTGCTTNTMGNGA 863
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
244  uAspProGlyAlaHisArgAsnLeuTyrPheHisIleAlaArgLeuArgVal 261
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
864  RGAYCCGCGCGCNCAYMGNAAYTNTGCAATATHCNMGNTNMGNGTNY 913
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
261  euserProGlyValTyrGlnLeuAspAlaProCysCysLeuProGlyLys 277
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
914  TWSNCCGCGNGTNTGCGCARTNGAYGCGCCNGCTGYTGYTTCNCGNAR 963
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
278  ValThrLeuCysThrProGlnAlaProAspGlnSerProCysGlnProLeuVal 294
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
964  GTNACNTNTGTYTGGCARGCNCNGAYCARMSNCCNTGYTCARCCNTYTGT 1013
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
294  IProProValProGlnLysAsnAlaThrValAsnGluProGlnAspPheG 311
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1014  NCNCCGCTGCCARARARAYGCGNACNGTAAAGARCCNARGATYTC 1063
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
311  LLeuValAlaGlyHisProAsnLeuCysValGlnValSerThrTyrGlu 327
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1064  ARTTNGTNGCGNCAYCCNAAAYTNTGYTNCARGTWSNACNTGGGAR 1113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
328  LysValGlnLeuGlnAlaCysSerThrAlaAspSerLeuGlyProPheLys 344
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1114  AARGTTCARTTNCARGCNTGYTNTGGCGNGAYWSNNTNGNCCNTTYAA 1163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
344  sAspAspMetLeuLeuValGluMetLysThrGlyLeuAsnAsnThrSerV 361
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1164  RGAYGAYATGYTNTNTNGARATGARACGNGNTTNAAYATACMSNG 1213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361  aLcysAlaLeuGluProSerGlyCysThrProLeuProSerMetAlaSer 377
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1214  TMTGYCANTNGARCCNMSNGNTGYACNCCNTNCCMWSNATGCGMSN 1263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
378  ThrArgAlaAlaArgLeuGlyGluGlnLeuLeuGlnAspPheArgSerH 394
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1264  ACNMGCGCGCCNMGNTNGSNGARGARTTNTNCARGATYTTTNGMSNCA 1313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
394  sGlnCysMetGlnLeuTyrAsnAspAspAspMetGlySerLeuTyrPal 411
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1314  YCARGTATATGCAATYTNCGAAYGAYATATGGGMSNYTNTGGCGNT 1363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
411  yspPrometAspLysTyrIleHis 418
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1364  GYCENATGCAATATATATCAAT 1386
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA001A.DAT: AAC85027

seq_documentation_block:

ID AAC85027 standard; cDNA; 2255 BP.

AC AAC85027;

DT 08-MAY-2001 (first entry)

DE Human cytokine receptor zcytor1a encoding cDNA.

XX

100

354 RGLYLeuAsnThrSerValCysAlaLeuGluProSerGlyCysThrP 371


```

1314 CCCCCAGACACAGATCCCTCTGTCCTTGGAACCCAGTGGCTGACTT 1363
371 roleProSerMetAlaSerThrArgAlaAlaArgLeuGluGluLeu 387
1364 CACTACCCAGCAAGCCTCTCAGAGGCGAGCTGCGCTTGAGAGTACTTA 1413
388 LeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsnAspAsp 404
1414 CTACAGAGACCTCAGTACAGGCGAGTCTGCAGACTATGG...GACGATGA 1460
404 nMetCysLeuLeuTrpAlaCysProMetAspLysTrpLeuHisArgArg 420
1461 CTGGGAGCGCTATGGCGCTCCGCCATGACAAATATCATCCAGACGCC 1509

eq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAS46223
eq_documentation_block:
D AAS46223 standard; cDNA; 2380 BP.
X
X AAS46223:
X
X 18-DEC-2001 (first entry)
X
X Human DNA encoding PRO polypeptide sequence #299.
X
X PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; sw;
X dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
X blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
X adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
X PCR primer.
X
X Homo sapiens.
X
X WO200168848-A2.
X
X 20-SEP-2001.
X
X
X 28-FEB-2001; 2001MO-US06520.
X
X
X 01-MAR-2000; 2000MO-US05861.
X 02-MAR-2000; 2000MO-US05861.
X 03-MAR-2000; 2000US-187202P.
X 06-MAR-2000; 2000US-186968P.
X 14-MAR-2000; 2000US-189320P.
X 14-MAR-2000; 2000US-189328P.
X 15-MAR-2000; 2000MO-US06884.
X 21-MAR-2000; 2000US-190828P.
X 21-MAR-2000; 2000US-191007P.
X 21-MAR-2000; 2000US-191048P.
X 21-MAR-2000; 2000US-191314P.
X 28-MAR-2000; 2000US-192655P.
X 29-MAR-2000; 2000US-193032P.
X 30-MAR-2000; 2000US-193053P.
X 30-MAR-2000; 2000MO-US058439.
X 04-APR-2000; 2000US-194447P.
X 04-APR-2000; 2000US-195975P.
X 11-APR-2000; 2000US-196000P.
X 11-APR-2000; 2000US-196187P.
X 11-APR-2000; 2000US-196690P.
X 11-APR-2000; 2000US-196820P.
X 18-APR-2000; 2000US-198121P.
X 18-APR-2000; 2000US-198585P.
X 25-APR-2000; 2000US-199397P.
X 25-APR-2000; 2000US-199550P.
X 25-APR-2000; 2000US-199554P.
X 03-MAY-2000; 2000US-201516P.
X 17-MAY-2000; 2000MO-US13705.
X 22-MAY-2000; 2000MO-US14042.
X 30-MAY-2000; 2000MO-US14941.
X 02-JUN-2000; 2000MO-US15264.
X 05-JUN-2000; 2000US-209832P.

```

```

PR 28-JUL-2000; 2000MO-US20710.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000MO-US23328.
PR 08-NOV-2000; 2000MO-US30952.
PR 01-DEC-2000; 2000MO-US32678.
PR 20-DEC-2000; 2000MO-US34956.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL,
XX Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2001-602746/68.
XX P-PSDB; AAU29322.
XX
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
XX presence of tumours, such as prostate and breast tumours, in mammals and
XX to screen for modulators of the compounds -
XX
XX Claim 2; Fig 597; 774pp; English.
XX
XX Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR
XX primers for PRO polypeptides of the invention. The sequences of the
XX invention can be used to detect the presence of a tumour in a mammal by
XX comparing the level of expression of a PRO polypeptide in a test sample
XX of cells from the animal and a control sample of normal cells, whereby a
XX higher level of expression in the test sample indicates the presence of a
XX tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,
XX pigs, goats and rabbits but are preferably human. The polypeptides can be
XX used to stimulate tumour necrosis factor (TNF) alpha release from human
XX blood, when contacted with it. A specific polypeptide can be used to
XX stimulate the proliferation or differentiation of chondrocyte cells. The
XX PRO proteins can be used to determine the presence of tumours and also
XX susceptibility to tumour development, particularly adrenal, lung, colon,
XX breast, prostate, rectal, cervical, or liver tumours, in mammalian
XX subjects. The oligonucleotide probes specific for the PRO nucleic acids
XX can be used for genetic analysis of individuals with genetic disorders.
XX
XX Sequence 2380 BP; 411 A; 776 C; 743 G; 450 T; 0 other;
XX
XX
XX alignment_scores:
XX      Quality: 1554.00      Length: 433
XX      Ratio: 4.281          Gaps: 3
XX      Percent Similarity: 83.834      Percent Identity: 67.667
XX
XX alignment_block:
XX US-09-899-471-2_COPY_21_440 x AAS46223 ..
XX
XX Align seg 1/1 to: AAS46223 from: 1 to: 2380
XX
XX 1 LeuGluArgLeuMetLupProGlnAspThrAlaArgCysSerLeuGlyLe 17
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 293 CTGGAGAGCGCTGTGGCGCTCAGAGCGTACCCAGCGCTCTCCGGCGCT 342
XX 17 USerCysHisLeuTrpAspGlnAspValLeuCysLeuProGlySerLeu 34
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 343 CTCTCTCGCGCTGTGGAGACAGTACACTCTGTGCTGGGAGACTCG 392
XX
XX 34 InsertArgGlyProValLeuValProThrArgLeuGlnThrGluLeu 50
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 393 TGCCTGTCTCCGGCGCGCTGTCTGGCGCTTACGACCTGCAGACAGACTG 442
XX
XX 51 ValLeuArgCysProGlnLysTrpAspCysAlaLeuArgValArgVal 67
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 443 GTGCTAGTGTCCAGAGAGAGACGACTGTGACTCTGTCTGCTGTGGC 492
XX
XX 67 ValHisLeuAlaValHisGlyHisTrpAlaGluProGluGlu..... 81
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 493 TGTCCACTTGGCGCTGATGGGCACTGGGAAGCCCTGAAGATGAGGAA 542
XX
XX 82 .....AlaGlyLysSerAspSerGluLeuGlnGluSerArgAsnAlaSer 96
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

543 AGTTTGAGAGACAGCTGACTCAGGGGTGAGAGAGCTAGCAATGCTCT 592
97  LeuGlnAlaGlnValValLeuSerPheGlnAlaTyrProIleAlaArgCys 113
   |||||.....|||  |||||.....|||
593 CTCAGAGCCCAAGTGTCTCTCTCCCTTCAGAGCTACCTACCTCCGCTG 642
113  sAlaLeuLeuGlnValGlnValProAlaAspLeuValGlnProGlyGlns 130
   |||||.....|||  |||||.....|||
643 CGTCTGTGTGAGGTGCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 692
130  erValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGln 146
   |||||.....|||  |||||.....|||
693 CTGTGGGCTGTGTGTATATAGTCTGTGTGAGGCTCTCTCTCTCTCTCT 742
147  ValGlnIleTyrPserTyrTyrTyrTyrTyrTyrTyrTyrTyrTyr 163
   |||||.....|||  |||||.....|||
743 GTAGCAATCTGTCTCTATATCTAGCCAGCTAGCAAGAGCACTCAACA 792
163  uThrGlnLeuPro.....AspGlyA 171
   |||||.....|||
793 CACACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 842
171  sPasnValLeuLeuThrLeuAspValSerGlnGlnGlnAspPheSerPhe 187
   |||||.....|||  |||||.....|||
843 ACAAGCTCATCTGTCTGTGATGTCTGTGAGAGAGCACTTCTGCTCTC 892
188  LeuLeuTyrLeuArgProValProAspAlaLeuLysSerLeuTyrTyr 204
   |||||.....|||  |||||.....|||
893 TCCCTGTACTGTAATCAGGTGCTCAGGGCCCCCAAAACCCGGTGCA 942
204  sAsnLeuThrGlyProGlnAsnIleThrLeuAsnIleThrAspLeuValP 221
   |||||.....|||  |||||.....|||
943 AAACCTGCTGTGAGCGCAGATCATCTTGAACACACAGACCTGCTGCT 992
221  rOCysLeuCysIleGlnValTyrPserLeuGlnProAspSerGlnArgVal 237
   |||||.....|||  |||||.....|||
993 CCTGCTGTGTATTCAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1042
238  GluPheCysProPheArgGlnAspProGlyAlaHisArgAsnLeuTyrP 254
   |||||.....|||  |||||.....|||
1043 AACATCTGCTCTCTCAGGAGAGACCCCGCACACAGAACCTCTGCGCA 1092
254  sIleAlaArgLeuArgValLeuSerProGlyValTyrPginLeuAspAlaP 271
   |||||.....|||  |||||.....|||
1093 AGCGGCCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1142
271  rOCysCysLeuProGlyLysValThrLeuCysTyrPginAlaProAspGln 287
   |||||.....|||  |||||.....|||
1143 CGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1192
288  SerProCysGlnProLeuValProProValProGlnLysAsnAlaThrVa 304
   |||||.....|||  |||||.....|||
1193 GACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1242
304  lAsnGlnProGlnAspPheGlnLeuValAlaGlnHisProAsnLeuCysV 321
   |||||.....|||  |||||.....|||
1243 GGACAGAGTTCGAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1292
321  aGlnValSerThrTyrPginLysValGlnLeuGlnAlaCysSerTyrPala 337
   |||||.....|||  |||||.....|||
1293 TTCAGGTGAAAGCTGCGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 1342
338  AspSerLeuGlyProPheLysAspPheLeuLeuValGlnLeuTyrThn 354
   |||||.....|||  |||||.....|||
1343 GACGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392
354  rGlyLeuAsnAsnThrSerValCysAlaLeuGlnProSerGlyCysThrP 371
   |||||.....|||  |||||.....|||
1393 CCCCAGAGACAAGATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1442
371  rOleuProSerMetAlaSerThrArgAlaAlaArgLeuGlyGlnLeu 387
   |||||.....|||  |||||.....|||
1443 CACTAACCCAGCAAAAGCTTCACAGAGGCTGCTGCTGCTGCTGCTGCT 1492

```

```

388  LeuGlnAspPheArgSerHisGlnCysMetGlnLeuThrPasnAspAspAs 404
   |||||.....|||  |||||.....|||
1493 CTACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1539
404  nMetGlySerLeuThrAlaCysProMetAspLysTyrIleHisArgArg 420
   |||||.....|||  |||||.....|||
1540 CTGGAGAGCGCTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1588
seq_name: /SIDS1/gcdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA509515
seq_documentation_block:
ID  AA509515 standard; cDNA, 2380 BP.
AC  AA509515;
XX
XX  24-Oct-2001 (first entry)
DE  Human cDNA encoding Interleukin 17 receptor, IL-17RH2.
XX
XX  Human; Interleukin-17 receptor; IL-17RH2; agonist; antagonist; ss;
KW  PRO20040; DNA 164625-2890; systemic lupus erythematosus;
KW  rheumatoid arthritis; osteoarthritis; diabetes mellitus;
KW  allergic disease; asthma; demyelinating disease;
KW  degenerative cartilaginous disorder; transplantation associated disease.
XX
OS  Homo sapiens.
XX
FH  Key Location/Qualifiers
FT  CDS 233..2347
   /*tag= a
   /*product= "IL-17RH2"
FT  sig_peptide 233..292
   /*tag= b
FT  mat_peptide 293..2344
   /*tag= c
   /*label= Mature_IL_17RH2
FT
FT
PN  WO200146420-A2.
XX
PD  28-JUN-2001.
XX
PF  20-DEC-2000; 2000WO-US34956.
XX
PR  23-DEC-1999; 99US-0172096.
PR  30-DEC-1999; 99WO-US31274.
PR  11-JAN-2000; 2000US-0175481.
PR  18-FEB-2000; 2000WO-US04341.
PR  02-MAR-2000; 2000WO-US05841.
PR  21-MAR-2000; 2000US-0191007.
PR  21-MAR-2000; 2000WO-US07532.
PR  02-JUN-2000; 2000WO-US15264.
PR  22-JUN-2000; 2000US-0213087.
PR  22-AUG-2000; 2000US-0644848.
PR  24-AUG-2000; 2000WO-US33328.
PR  24-OCT-2000; 2000US-0242837.
PR  10-NOV-2000; 2000WO-US30873.
PR  28-NOV-2000; 2000US-0253646.
PR  01-DEC-2000; 2000WO-US32678.
XX
PA  (GENTECH ) GENENTECH INC.
XX
XX  Chen J, Filvaroff E, Fong S, Goddard A, Godowski PJ, Grimaldi CJ,
PI  Gurney AL, Li H, Hillan KJ, Tamas D, Van Lookeren M, Vandlen RJ,
PI  Watanabe CK, Williams PM, Wood WI, Yansura DG;
DR  WPI: 2001-451708/48.
XX  P-PSDB: AA004956.
XX
XX  Novel PRO polypeptides homologous to interleukin-17, useful for the
PT  diagnosis and treatment of immune related disease e.g. rheumatoid
PT  arthritis and diabetes -
XX

```


PS Claim 1, Fig 13; 188bp; English.

CC The sequence (DNA 164625-2890) encodes a PRO polypeptide (PRO20040)
 CC which is the human interleukin 17 receptor, IL-17R2. A composition
 CC containing ant/agonists to the PRO polypeptides or individual components
 CC are useful for treating a mammal with an immune related disease, e.g.
 CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, an
 CC juvenile chronic arthritis, a spondyloarthritis, systemic sclerosis, an
 CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
 CC disease, a demyelinating disease, an autoimmune or immune-mediated skin
 CC disease, contact dermatitis, an allergic disease e.g. food
 CC hypersensitivity, asthma, a transplantation associated disease, or a
 CC chronic inflammatory demyelinating polyneuropathy. Treating a
 CC degenerative cartilaginous disorder comprises administering a PRO1031 or
 CC PRO1122 polypeptide agonist, or antagonist to the mammal. Numerous
 CC examples of the diseases and disorders are given in the specification.
 CC
 CC Sequence 2380 BP; 411 A; 776 C; 743 G; 450 T; 0 other;

Alignment_scores:
 Quality: 1554.00 Length: 433
 Ratio: 4.281 Gaps: 3
 Percent Similarity: 83.834 Percent Identity: 67.667

Alignment_block:
 US-09-899-471-2_COPY_21_440 x AAS09515 ..

Align seg 1/1 to: AAS09515 from: 1 to: 2380

```

1 LeuGIuArgLeuMeGclupProGlnAAspThrAlaArgCysSerLeuGlyLe 17
23 CTGGAGAGAGCTTGTGGGGCTCAGAGACGCTACCCACCTGCTCCGGGCT 342
17 uSerCysHisLeuTrpAspGlyAspValLeuCysLeuProGlySerLeuG 34
343 CTCCTGGCGGCTCTGGGACAGTACATCTGCTGCTGGGACATCG 392
34 InSerAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeu 50
393 TGCCTGCTCCGGGCGGCTGCTGGCGCTGACGACCTGCTGCTGCTGCTG 442
51 ValLeuArgCysProGlnLysThrAspCysAlaLeuArgValArgVala 67
443 GTGCTGAGGTGCGCAGAGAGACGCTGACCTGCTGCTGCTGCTGCTGCTG 492
67 lValHisLeuAlaValHisGlyHisTrpAlaGluProGlu..... 81
493 TGTCCACTGGCGGCTGATGGGCACTGGGAGAGCGTAGAGATGAGGAA 542
82 .....AlaGlyLysSerAspSerGluLeuGlnLysArgAsnAlaSer 96
543 AGTTGGAGGACGACTGACTGAGGGGTGAGAGAGCTTAGAATGCTCT 592
97 LeuGlnAlaGlnValValLeuSerPheGlnAlaTyrProTleAlaArgC 113
593 CTCACGGCCCAAGTGTGCTGCTCTCCAGGCTTACCTTACCTGCGCGCTG 642
113 sAlaLeuLeuGlnValGlnValProAlaAspLeuValGlnProGlyGln 130
643 CGTCCCTGCTGAGGTGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 692
130 erValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGln 146
693 CTGTGGGCTCTGTGTATATGACTGCTTGGAGGCTCCCTAGGAGAGTAG 742
147 ValGlnIleTyrSerTyrThrLysProArgTyrGlnLysGlnLeuAsn 163
743 GTAGCAATCTGCTCTATCTACAGCCAGGTACGAGAGAGAACTCAACCA 792
163 uThrGlnLysLeuPro.....AspGlyA 171

```

```

793 CACACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 842
171 sPsnValLeuLeuThrLeuAspValSerGluGlnAlaPheSerPhe 187
843 ACACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 892
188 LeuLeuTyrLeuArgProValProAspAlaLeuLysSerLeuTyrTyr 204
893 TCCTGTACAGCAATGATGATGATGATGATGATGATGATGATGATGATG 942
204 sAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAspLeuVal 221
943 AACCTGACCTGAGACCGCAGATCATTCCTGATACACACACAGCTGTTC 992
221 rocYsCysLysIleGlnValTyrPheSerLeuGlnProAspSerGlnArg 237
993 CTTGCTCTTATTCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1042
238 GluPheCysProPheArgGluAspProGlyValAlaHisArgAsnLeuTrp 254
1043 AACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1092
254 sIleAlaArgLeuArgValLeuSerProGlyValTyrPheLeuAspAla 271
1093 AGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1142
271 rocYsCysLeuProGlyLysValThrLeuCysTyrPheAlaProAspGln 287
1143 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1192
288 SerProCysGlnProLeuValProProValProGlnLysAsnAlaThrVa 304
1193 GACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1242
304 lAsnGluProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuCysV 321
1243 GGAACAGGCTTCTGAGTTCATGCTGTAAGGCCACCTTAACCTCTG 1292
321 alGlnValSerThrTyrPheGluLysValGlnLeuGlnAlaCysSerTrpAla 337
1293 TTCAGGTGAACAGCTCGAGAAAGCTGCAGAGAGTCTTGGGCT 1342
338 AspSerLeuGlyProPheLysAspAspMetLeuLeuValGluMetLysTh 354
1343 GACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392
354 rGlyLeuAsnAsnThrSerValCysAlaLeuGluProSerGlyCysTrp 371
1393 CCCCAGAGAACAGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1442
371 roleuProSerMetAlaSerThrArgAlaAlaArgLeuGlyGlnLeu 387
1443 CACTACCCACCAAGCCCTCCACGAGGCGAGCTCGCTGGAGAGTACTTA 1492
388 LeuGlnAspPheArgSerHisGlnCysMetClnLeuTrpAsnAspAsp 404
1493 CTACAGAGCTGCTGAGTACAGGCACTGCTGCTGCTGCTGCTGCTGCTG 1539
404 nMetGlySerLeuTrpAlaCysProMetAspLysTyrIleHisArgArg 420
1540 CTTGGGAGCGCTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1588
seq_name: /SIDS1/gcgdata/geneseq/geneseg-emb1/NA2001A.DAT:AAF92138
seq_documentation_block:
ID AAF92138 standard; cDNA; 2380 BP.
AA92138:
AC AAF92138:
XX
XX
DT 15-MAY-2001 (first entry)
XX
DE Human PRO20040 cDNA.

```



```

51 ValLeuArgCysProGlnIysThrAspCysAlaLeuArgValValVal 67
443 GTGCTGAGCTGCCAGAGAGACCGACTGTGACTCTGTCTGCGCTGGC 492
67 IValHisLeuAlaValHisGlyHisIleThrAlaGluProGluGlu 81
493 TGTCCACTGTGGCCGTGCATAGGGGACCTGGGAAGCCTGMAATGGAGAA 542
82 ....AlaGlyLysSerAspSerGluLeuGlnGluSerArgAsnAlaSer 96
543 AGTTTGGAGGAGACACTGTACTAGGGGTGGAGGACCTTAGAATGCCCT 592
97 LeuGlnAlaGlnValValLeuSerPheGlnAlaIleArgProIleAlaArgCys 113
593 CTCGAGGCCCAAGTGTGTCTCTCTCCAGGCTACCTACTGCCCCGTG 642
113 sAlaLeuLeuGlnValValGlnValProAlaAspLeuValGlnProGlyGlns 130
643 CGTCTGCTGTGAGGTGAAGTGCCTCTGCTGCTTGTGCAGTTGGTCAGT 692
130 eValGlySerAlaAlaPheAspCysPheGluAlaSerLeuGlyAlaGlu 146
693 CTGGGGCTGTGGTAAATGACTGCTTGAGGCTGCCCTTAGGAGAGTAG 742
147 ValGlnIleIlePheSerIleThrLysProArgIleGlnIleGlnAspLe 163
743 GTAGCAATCTGTGCTTAACTAATCTAGCCCAAGTACGAGAAGAACTCAACA 792
163 uThrGlnGlnLeuPro.....AspGly 171
793 CACACACAGCTGCTGCCCTGCCCTGGGCTCAACGTTCAAGCAATGCTG 842
171 spAsnValLeuLeuThrLeuAspValSerGluGlnGlnAspPheSerPhe 187
843 ACAACGGTCATCTGGTTGTGAATGTCTTGAGAGAGACACTTGGGCTC 892
188 LeuLeuArgIleuArgProValProAspAlaLeuLysSerLeuIlePyl 204
893 TCCTGTACTGTGAATCAGTCCAGGGCCCCCAAAACCCGGTGGCACAA 942
204 sAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisIleThrAspLeuValP 221
943 AAACCTCACTGGACCGCAGATCATTAACCTTGAACCAACAGACTGGTTC 992
221 rocYsLeuCysIleGlnValIlePheSerLeuGluProAspSerGluArgVal 237
993 CTGCTCTGTATTCAAGTGTGGCTCTGGAACTGCTCGTTAGGAGCG 1042
238 GluPheCysProPheArgGlnAspProGlyAlaHisArgAsnLeuThrP 254
1043 AACATCTGCCCTTCACAGGAGAGACCCCGCGCACACCAAGAACTCTGGCA 1092
254 sIleAlaArgLeuArgValLeuSerProGlyValIlePheGlnLeuAspAlaP 271
1093 AGCGCCCGCAGCTGAGACTGTGACCTCGACAGAGCTGGCTGTGACCCAC 1142
271 rocYsCysLeuProGlyLysValIleThrLeuCysThrGlnAlaProAspGln 287
1143 CGTGTCTGCTGCCCGCAGAAACGGCACTGTCTGCGGGCTCCGGGGTGG 1192
288 SerProCysGlnProLeuValProProValProGlnIleAsnAlaThrVa 304
1193 GACCCCTGCCAGCCACTGTGTCCACCGGCTTCTCGGAGAAACCTCATCTGT 1242
304 lAsnGluProGlnAspPheGlnLeuValAlaGlnHisProAsnLeuCysV 321
1243 GGCACAAGTCTCGAGTTCATCTGCTAAAGGCCACCTTAACCTCTGTG 1292
321 aGlnValSerThrIlePheGluLysValGlnLeuGlnAlaCysSerThrPala 337
1293 TTCAGGTGAACAGCTCGAGAAAGCTGCAGCTCAGAGAGTGTGTGGCT 1342

```


338 AspSerLeuGlyProPheLysAspAspMetLeuValGluMetLysTh 354
 1343 GACTCCCTGGGGCTCTCAAGACATGCTACTGTTGAGACACAGAGC 1392
 354 rGlyLeuAsnThrSerValLysAlaLeuGluProSerGlyCysThp 371
 1393 CCCCCAGACAAAGATCCCTCTGCTTGGAAACCCAGTGGCTACTT 1442
 371 rGlyProSerMetAlaSerThrArgAlaAlaArgLeuGlyGluLeu 387
 1443 CACTACCCACCAAGCCTCCACGAGGAGAGTCCCTTGGAAAGACTTA 1492
 388 LeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsnAspAsp 404
 1493 CTCAAGACCTGACAGTCCAGCCAGTGTCTGACGATATG...GACGATGA 1539
 404 nMetGlySerLeuTrpAlaCysProMetAspLysTrpIleHisArg 420
 1540 CTGGGAGGCGCTATGGGCTGCCCCCATGACAAATACATCCCAAGCCG 1588
 seq_name: /SID1/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT.AA265269
 seq_documentation_block:
 AA265269 standard; DNA; 1752 BP.
 AA265269;
 23-MAR-2000 (first entry)
 Human secreted protein gene 20.
 Human; secreted protein; cancer; tumour; developmental abnormality;
 foetal deficiency; blood disorder; immune system disorder; inflammation;
 autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
 schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
 atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
 digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
 therapy; chromosome 3; ds.
 Homo sapiens.
 WO958660-A1.
 18-NOV-1999.
 06-MAY-1999; 99MO-US09847.
 12-MAY-1998; 98US-0085093.
 12-MAY-1998; 98US-0085094.
 12-MAY-1998; 98US-0085105.
 12-MAY-1998; 98US-0085180.
 18-MAY-1998; 98US-0085906.
 18-MAY-1998; 98US-0085920.
 18-MAY-1998; 98US-0085921.
 18-MAY-1998; 98US-0085922.
 18-MAY-1998; 98US-0085923.
 18-MAY-1998; 98US-0085924.
 18-MAY-1998; 98US-0085928.
 18-MAY-1998; 98US-0085925.
 18-MAY-1998; 98US-0085927.
 (HUMA-) HUMAN GENOME SCI INC.
 Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
 Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;
 Latleur DW, Endress GA, Edner R;
 MPI; 2000-062296/05.
 P-PSDB; AAY76143.
 New isolated human genes and the secreted polypeptides they encode,
 useful for diagnosis and treatment of e.g. cancers, neurological
 disorders, immune diseases, inflammation or blood disorders

xx Claim 1; Page 308; 475pp; English.
 ps This sequence was found to be present on human chromosome 3.
 xx AA265269 to AA265350 represent 97 isolated human secreted protein genes.
 CC This sequence was found to be present on human chromosome 3.
 CC AA265269 to AA265350 represent 97 isolated human secreted protein genes.
 CC The genes and their corresponding secreted polypeptides are
 CC useful for preventing, treating or ameliorating medical conditions,
 CC e.g. by protein or gene therapy. Also pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new genes. Specific
 CC uses are described for each of the 97 genes, based on which tissues they
 CC are most highly expressed in, and include developing products for the
 CC diagnosis or treatment of cancer, tumours, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
 CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
 CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
 CC disorders, digestive/endocrine disorders, infections and AIDS. The
 CC polypeptides are also useful for identifying their binding partners.
 CC The sequences shown in AA265269 to AA265350 represent fragments of the
 CC secreted proteins.
 xx Sequence 1752 BP; 324 A; 555 C; 522 G; 347 T; 4 other;
 xx

alignment_scores:
 Quality: 1505.50 Length: 448
 Ratio: 4.205 Gaps: 3
 Percent Similarity: 79.911 Percent Identity: 64.732

alignment_block:

US-09-899-471-2_COPY_21_440 x AA265269

Align seg 1/1 to: AA265269 from: 1 to: 1752

1 LeuGluArgLeuMetGluProGlnAspThrAlaArgCysSerLeuGlyLe 17
 287 CTGGAGAGGCTGTGGGCGCTCAGACGACGACGACGACGACGACG 336
 17 uSerCysHisLeuTrpAspGlyAspValLeuCysLeuProGlySerLeu 34
 337 CTCCCTGCCCGCTCTGGGACATGACATGCTCTGCTGGGACATG 386
 34 InsrAlaProGlyProValLeuValProThrArgLeuGlnInrGluLeu 50
 387 TGCTGCTCCGGGCCCCGCTGCTGGCCCTACGACGACGACGACGAC 436
 51 ValLeuArgCysProGlnLysThrAspCysAlaLeuArgValAla 67
 437 GTGCGAGGTGCGAAGAGAGAGACGACTGACCTGCTGCTGCTGCG 486
 67 lValHisLeuAlaValHisGlyHisTrpAlaGluProGluGlu..... 81
 487 TGTCCACTTGGCGCTGCATGCGCCTGGAAGAGCCGGAAGATGAGAA 536
 82AlaGlyLysSerAspSerGluLeuGlnGlnSerArgGlnAlaSer 96
 537 AGTTGGAGGACGACGCTGACTTATAGGGGTGAGAGGCTTACGAATG 586
 97 LeuGlnAlaGlnValValLeuSerPheGlnAlaTyrProPheAlaArg 113
 587 CTCACGCCCCAAGTCTGCTCTCTCTCCAGCCTACCTACTGCCCCCT 636
 113 SalAlaLeuGluValGlnValProAlaAspLeuValGlnProGlyLys 130
 637 GCTCTGCTGAGGTGCAAGTCCCTGCTGCTGCTGCTGCTGCTGCT 686
 130 eValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGlu 146
 687 CTGTGGCTCTGTGGTATATGACTGCTGAGGCTGCTGCTGAGGAGTG 736
 147 ValGlnIleTrpSerTyrThrLysProArgTyrGlnLysGluLeuAsn 163

[illegible][illegible]

291 CTCCTCCGCCCTCTGGGACATGACATCTCTGCTGCTGGGACATCG 340
 34 InsertAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeu 50
 341 TGCCCTGCTCGGGGCCCTGCTGCTGGCCCTACGACATCGACAGACAGCTG 390
 51 ValLeuArgCysProGlnIlyThrAspCysAlaLeuArgValArgVala 67
 391 GTCTGAGAGTCCAGAGAGACAGACCTGACCTGCTGCTGCTGCTGCTGCTG 440
 67 ValHisLeuAlaValHisGlyHisThrAlaGluProGluGlu..... 81
 441 TGTCCACTGGCCGCTGATGAGCAGCTGGAGAGCCGTGAAGATGAGGAAA 490
 82AlaGlySerSerAspSerGluLeuGlnGlnIlySerArgAsnAlaSer 96
 491 AGTTTGAGAGACAGCTGACTTGGGGTGGAGAGCCTAGCAATGCTCT 540
 97 LeuGlnAlaGluValValLeuSerPheGlnAlaThrProIleAlaArgCys 113
 541 CTCGAGGCCCAAGTCTGCTCTCTCCACAGCCCTACCTACTGCTGCTGCTG 590
 113 sAlaLeuGluValGluValProAlaAspLeuValGluProGluGln 130
 591 CGTCTGCTGGAGGTGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
 130 erValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGlu 146
 641 CTGTGGGCTCTGTGATATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 690
 147 ValGlnIleThrSerThrThrIlyProArgTyrlGlnIlyGluLeuAsnLe 163
 691 GTACCAATCTGCTCTATCTACTACAGCCAGGTAGAGAAAGCACTCAACCA 740
 163 uThrGlnIleuPro..... 168
 741 CACACAGCAGCTGCTGACTGACAGGGGGCTGCAAGTCTGCAACAGCATCC 790
 169AspGlyAspAsn 172
 791 CGAGCTGTGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 173 ValLeuLeuThrLeuAspValSerGluGlnGlnAspPheSerPheLeuLe 189
 841 GTGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 890
 189 uTyrlLeuArgProValProAspAlaLeuIlySerLeuThrIlyTrpIlyAsnL 206
 891 GTACTGGAATCAAGTCCAGGGCCCCCAAAACCCGGTGGCACAACCAACCC 940
 206 euThrGlyProGlnAsnIleThrLeuAsnHisThrAspLeuValProCys 222
 941 TGACGTGAGACCGAGATCACTTACCTGACACACAGACCTGCTGCTGCTG 990
 223 LeuGlyIleGlnValIlyPheSerLeuGluProAspSerGluArgValGluPh 239
 991 CTCTGTATTCAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1040
 239 eCysProPheArgGluAspProGlyAlaHisArgAsnLeuThrIleHisLea 256
 1041 CTGCCCCCTTCAGGGAGACCCCGCCGACACCAACCTCTGGGAAAGCCG 1090
 256 laArgLeuArgValLeuSerProGlyValIlyPheGlnLeuAspAlaProCys 272
 1091 CCCGACTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 273 CysLeuProGlyIlyValThrLeuGlyTrpGlnAlaProAspGlnSerPr 289
 1141 TCGCTGCTCCCGCAAGACCGGACACTGTCTGCTGCTGCTGCTGCTGCTGCTG 1190
 289 oCysGlnProLeuValProProValProGlnIlyAsnAlaThrValAsnG 306
 1191 CTGCGACGACCATGCTCCACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1239

306 IupProGlnAspPheGlnLeuValAlaGlnHisProAsnLeuGlyValGln 322
 1239 1239
 323 ValSerThrTrpGluIlyValGlnLeuGlnAlaCysSerThrPalaspse 339
 1240 GTCAACAGCTGCTGAGAGAGCTGACCTGACAGAGTCTGTGTGGCTGACTC 1289
 339 rLeuGlyProPheIlyAspAspMetLeuValGlnLeuIlyThrGlyL 356
 1290 CTTGGGCTCTCTCAAGACAGTGTCTACTGTGTGAGACACAGGCCCCCC 1339
 356 euAsnAsnThrSerValCysAlaLeuGluProSerGlyCysThrProLeu 372
 1340 AGGACAAACAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1389
 373 ProSerMetAlaSerThrArgAlaAlaArgLeuGlyGluLeuLeuGlu 389
 1390 CCCAGCAACCTCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1439
 389 naSpPheArgSerHisGlnCysMetGlnLeuThrAsnAspAsnMetG 406
 1440 AGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1486
 406 IySerLeuThrPalCysProMetAspIlyTrpIleHisArgArg 420
 1487 GACGCGTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1530
 seq.name: /SID1/gc9data/geneseq/geneseqn-emb1/NA2001A.DAT:AAF93875
 seq.documentation_block:
 ID AAF93875 standard; cDNA; 2499 BP.
 AC AAF93875;
 DT 23-MAY-2001 (first entry)
 DE Human cDNA encoding a membrane or secretory protein clone PSEC0233.
 KW Human; secretory protein; membrane protein; vaccine; gene therapy;
 KW Rheumatoid arthritis; diabetes; ss.
 OS Homo sapiens.
 PN EP1067182-A2.
 XX 10-JAN-2001.
 PD 07-JUL-2000; 2000EP-0114090.
 PF 08-JUL-1999; 99JP-0194179.
 PR 11-JAN-2000; 2000JP-0118775.
 PR 02-MAY-2000; 2000JP-0183766.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Kawachi T, Sugiyama T, Hayashi K;
 DR WPI: 2001-093989/11.
 DR P-PSDB; AAB88448.
 XX Nucleic acids encoding secretory proteins/membrane proteins, useful in
 PT gene therapy or as candidate target molecules in drug development -
 XX
 PS Claim 1; SEQ ID 263; 609pp + CD ROM; English.
 XX
 CC This invention relates to nucleic acid sequences AAF93744 - AAF93916
 CC which encode human secretory or membrane proteins represented by
 CC AAB88317 - AAB88419. Included in the invention are primers
 CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
 CC cDNA sequences of the invention. The invention also includes methods for
 CC the production of antibodies directed against the proteins, and cDNA

CC sequences, which can be used in vaccines. The polynucleotide sequences
 CC can be used in gene therapy. The polynucleotide sequences and the
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with inappropriate secretory
 CC protein/membrane protein expression. The nucleic acids and complementary
 CC sequences may also be used as DNA probes in diagnostic assays
 CC (e.g. polymerase chain reactions (PCR)) to detect and quantify the
 CC presence of similar nucleic acid sequences in samples. They may also be
 CC used to study the expression and function of secretory proteins/membrane
 CC polypeptides and their role in metabolism. The polypeptides may be used
 CC as antigens in the production of antibodies against them and in assays to
 CC identify modulators (agonists and antagonists) of expression and
 CC activity. The antibodies and antagonists may also be used as therapeutic
 CC agents to down regulate expression and activity. The antibodies may also
 CC be used as diagnostic agents for detecting the presence of the
 CC polypeptides in samples (e.g. by enzyme linked immunosorbent assay
 CC (ELISA)). Examples of diseases which may be treated include rheumatoid
 CC arthritis and diabetes.

XX Sequence 2499 BP: 418 A; 821 C; 764 G; 496 T; 0 other;

alignment_scores:

Quality: 1397.00 Length: 504
 Ratio: 4.061 Gaps: 6
 Percent Similarity: 68.254 Percent Identity: 54.960

alignment_block:

US-09-899-471-2_COPY_21_440 x AAF93875

Align seg 1/1 to: AAF93875 from: 1 to: 2499

1 LeuGIuArgLeuMetGluProGlnAspThrAlaArgCysSerLeuGlyLe 17
 268 CTGAGAGGCTGTGGGCGCTCAGACGCTACCCAGCTCTCCGGGCT 317
 17 uSerCysHisLeuTrpAspGlyAspValLeuCysLeuProGlySerLeuG 34
 318 CTCCTGGCGGCTGTGGAGAGTGAATCTCTGCTGCTGGGAGACATCG 367
 34 InseuAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeu 50
 368 TGCCTGCTCCGGGCGGCGGCTGAGGAGCTGACCTGCTGCTGGGAGCTG 417
 51 ValLeuArgCysProGlnLysThrAspCysAlaLeuArgValAlaArgVal 67
 418 CTGCTGAGAGTGCACAGAGAGACCGACTGTGACTCTGCTGCTGGC 467
 67 ValHisLeuAlaValHisGlyHisTrpAlaGluProGluGlu..... 81
 468 TGTCCACTTGGCGGCTGACATGGGCACTGGGAGAGAGCTGAAGATGAGGAA 517
 82AlaGlyLysSerAspSerGluLeuGlnGluSerArgAsnAlaSer 96
 518 AGTTGGAGGAGCACTGACTGAGGGGTGAGAGAGCTGAGAAATGCTCT 567
 97 LeuGlnAlaGlnValValLeuSerPheGlnAlaTrpProIleAlaArgCys 113
 568 CTCAGAGCCCAAGTGTCTCTCTCTCCAGGCTTACCTACTGCTCCGCTG 617
 113 AlaLeuLeuGlnValValGlnValProAlaAspLeuValGlnProGlyGln 130
 618 CGTCTCTGTGAGGTGAAGTGCCTGTGCTGCTGTGCTGCTGCTGCTGCT 667
 130 erValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlnAlaGln 146
 668 CTGGGCTCTGTGTATATGATGCTCTGAGGCTGCTGCTGCTGCTGCTGCT 717
 147 ValGlnIleTrpSerTrpThrLysProArgTrpGlnLysGluLeuAsnIle 163
 718 GTAGAGATCTGTGCTTACTACAGCCCAAGTACGATGAGAACTCAACCA 767
 163 uThrGlnGlnLeuPro.....AspGlyA 171

768 CACACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1111111111
 171 spAsnValLeuLeuThrLeuAspValSerGluGluGlnAspPheSerPhe 187
 818 ACAACGTGATCTGTGTTCTGATATGCTCTGAGAGACAGACTTGGGCTC 867
 188 LeuLeuTrpLeuArgProValProAlaLeuLysSerLeuTrpTyrIly 204
 868 TCCCTGTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 917
 204 sAsn..... 205
 918 AAACCTGTGAGGCTTCCCTTCCCAAGTCCATCCCACTGTAGGCCGA 967
 205 205
 968 TGCCTGTGCAAGAGCAGCAGTCCATATCAGAGAGATCTTGAAGAGA 1017
 205 205
 1018 CTCACCCCAAGAGGAAATGTGTGGGAGACTTGTGCTTCTGCTGCT 1067
 205 205
 1068 TCCCTGACTTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1117
 206.....Leu ThrGlyProGlnAsnIleThrLeuAsnHisIsth 217
 1118 CCTTATTATTTCACACAGACTGGAGCCGAGATCATTTACTTGAACACAC 1167
 217 rAspLeuValProCysLeuCysIleGlnValTrpSerLeuGlnProAsp 234
 1168 AGACATAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1217
 234 erGluArgValGluPheCysProPheArgGlnAspProGlyAlaHisArg 250
 1218 CGTTTGGAGAGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267
 251 AsnLeuTrpHisIleAlaArgLeuArgValLeuSerProGlyValTrpG 267
 1268 AACCTGTGCAACCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1317
 267 nLeuAspAlaProCysCysLeuProGlyLysValThrLeuCysTrpGln 284
 1318 GCTGAGGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1367
 284 lArProAspLysTrpCysGlnProLeuValProProValProGlnLys 300
 1368 CTCGGGTGGGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1417
 301 AsnAlaThrValAsnGluProGlnAspPheGlnLeuValAlaGlyHisIsth 317
 1418 AACGTCACCTGTGACAGAGCTTCTGAGTTCCATGTGTTGAAGAGGCCACCC 1467
 317 AsnLeuCysValGlnValSerThrTrpGlnLysValGlnLeuGlnAla 334
 1468 TAACTCTGTGTACGTGAGTGAAGAGCTGAGAGAGCTGACCTGAGAGAGT 1517
 334 ySerTrpAlaAspSerLeuGlyProPheLysAspPheLeuVal 350
 1518 CTTTGTGGCTGACTCCCTGAGGCTCTCTCAAGACGATGTGCTACTGTTG 1567
 351 GluMetLysThrGlyLeuAsnAsnThrSerValCysAlaLeuGlnProse 367
 1568 GAGACAGCAGGCCCCCAGAGCAACATGCTTGTGCTTGTGAACCCAG 1617
 367 rGlyCysThrProLeuProSerMetAlaSerThrArgAlaAlaArgLeu 384
 1618 TGGCTGTACTTACTACACAGCAAAAGCTTCACAG..... 1651
 384 LysGluGlnLeuLeuLysAspPheArgSerHisGlnCysMetGlnLeuTrp 400
 11111111

1652CTATGG 1657
401 AsnAspAspAsnMetGlySerLeuTrpAlaCysProMetAspLysIle 417
1658 ...GACGATGACTTGAGAGCCCTATGGCTCCCTCCATGACAAATATCAT 1704
417 eHisArgArg 420
1705 CCACACAGCGC 1714

eq_name: /STDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: AAC85029

eq_documentation_block:

D AAC85029 standard; cDNA; 1753 BP.

X AAC85029;

X 08-MAY-2001 (first entry)

X Human variant Zcytor14 protein Zcytor14-1 encoding cDNA.

X Cytokine receptor: Zcytor14; human; inflammation; rheumatoid arthritis;

X antinflammatory; gene therapy; vaccine; variant; Zcytor14-1; ss.

X Homo sapiens.

X Key Location/Qualifiers

X CDS 2..1729

X /tag= a

X /product= "Zcytor14-1"

X WO200104304-A1.

X 18-JAN-2001.

X 30-JUN-2000; 2000MO-US18383.

X 07-JUL-1999; 99US-0348854.

X (ZYMO) ZYMOGENETICS INC.

X Presnell SR, Burkhead SK, Powder SL;

X WPI; 2001-112618/12.

X P-PSDB: AAB61881.

X New polypeptide encoding a human cytokine receptor Zcytor14, for
X treating inflammation e.g. rheumatoid arthritis -

X Disclosure: Page 95-98; 112pp; English.

X The invention provides a new human cytokine receptor designated Zcytor14.
X Zcytor14 can be expressed by standard recombinant methodology. The
X encoding nucleic acid is useful for detecting the expression of a
X Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be
X used to screen biological samples in vitro for the presence of Zcytor14.
X proteins, polypeptides and peptides having Zcytor14 activity can be
X administered to a subject who lacks an adequate amount of this
X polypeptide. For treating inflammation and conditions such as rheumatoid
X arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14
X antibodies) can be used to treat a subject who produces an excess of
X Zcytor14. Zcytor14 nucleotide sequences can also be used to provide
X Zcytor14 to a subject. The present sequence represents a cDNA encoding
X a variant of the human cytokine receptor Zcytor14, designated Zcytor14-1.
X The variant is a truncated form of the receptor polypeptide and lacks
X amino acid residues 1-113 of Zcytor14.

X Sequence 1753 BP; 287 A; 584 C; 548 G; 334 T; 0 other;

X Alignment_scores:

X Quality: 1154.50

X Ratio: 4.260

X Length: 340

X Gaps: 3

Percent Similarity: 79.706 Percent Identity: 65.588

alignment_block:

us-09-899-471-2_copy_21_440 x AAC85029 ..

Align seg 1/1 to: AAC85029 from: 1 to: 1753

90 GlnGlnSerArgAsnAlaSerLeuGlnAlaGlnValValLeuSerPheGln 106
:::|||||
2 GAGGAGCCTAGAAATGCTCTCTCCAGGCCCAAGTGTGCTCTCTCCCA 51
106 nAlaTyrProIleAlaArgCysAlaLeuLeuGlnValGlnValProIleA 123
|||||
52 GGCCTACCTACTGCCCGCTGCTCTGCGAGGAGTGCAGAGTGCCTCTG 101
123 sPLeuValGlnProGlnGlnSerValGlySerAlaValPheAspCysPhe 139
|||||
102 CCTGTGTCAGTTTGTGTCAGTCTGTGGGCTGTGTGATATGACTGCTTC 151
140 GlnAlaSerLeuGlnValGlnValGlnIleTrpSerTyrThrLysProAr 156
|||||
152 GAGGCTGCCCTAGGAGTGAAGTACGAATGTGCTATATCTAGCCAG 201
156 gTyrGlnLysGlnLeuAsnLeuThrGlnGlnLeuPro..... 168
202 GTACGAGAAGAACTCAACACACACAGCTGCTGCCCTGCCCTGCC 251
169AspGlyAspAsnValLeuLeuThrLeuAspValSer 180
252 TCACGGTGCACAGCAGATGGTGAACAGTGCATCTGTGATGATGCTCT 301
181 GlnGlnGlnAspPheSerPheLeuLeuTyrLeuArgProValProAspAl 197
|||||
302 GAGGAGCAGCAGCTTGGCTCTCTCTGACTGGAATGAGTCCAGGGCCC 351
197 aLeuLysSerLeuTyrPheLysAsnLeuThrGlyProGlnAsnIleThrL 214
|||
352 CCCAAACCCCGGTGGCACAAAACCTGACTGAGCCCAATATCATTACT 401
214 euAsnHisThrAspLeuValProCysLeuCysIleGlnValTrpSerLeu 230
|||||
402 TGAACACACAGACCTGTTCCCTGCTGTATTCAGTGTGGCTCTGTG 451
231 GlnProAspSerGlnArgValGlnPheCysProPheArgGlnAspProGln 247
|||||
452 GAACCTGACTCGGTAGAGACAGAACCTGCCCTTACGGAGAGACCCCG 501
247 yAlaHisArgAsnLeuTrpHisIleAlaArgLeuArgValLeuSerProG 264
|||||
502 CGCACACAGAAACCTTGGCAAGCCCGCAGCTGCCAGCTGACCTGC 551
264 lValAlaTrpGlnLeuAspAlaProCysCysLeuProGlnLysValThrLeu 280
|||
552 AGAGCTGGCTCTGGAGCAGCAGCTGCTGCTGCCGAGAGAGGGCAGCTG 601
281 CysTrpGlnAlaProAspGlnSerProCysGlnProLeuValProProVa 297
|||||
602 TGCTGCGCGGCTCGGGTGGGAGCCCTGCCACCTGTGTCTCCACCGCT 651
297 lProGlnLysAsnAlaThrValAsnGlnProGlnAspPheGlnLeuValA 314
:::|||||
652 TTCCTGGGAGAACGTCACCTGTGCAC..... 676
314 lAlaGlnHisProAsnLeuCysValGlnValSerThrTrpGlnLysValGln 330
|||||
677GTACACAGCTCGAGAGAGCTCAG 700
331 LeuGlnAlaCysSerTrpAlaAspSerLeuGlyProPheLysAspAspMe 347
|||||
701 CTGCAGAGATGCTGTGGGCTGACTCCCTGGGCGCTCTCAAGACGATGT 750
347 tLeuLeuValGlnMetLysThrGlyLeuAsnAspThrSerValCysAlaL 364
:|||||
:::|||||


```

169 .....AspGlyAspAsnVal 173
614 SNTGYTGCGCNYTNCNTGTYTNAAYGTWMSNGCAGVAGNGAYAGN 663
174 LeuLeuThrLeuAspValSerGluGluGlnAspPheSerPheLeuLeu 190
664 CATTNGTNTYNAAGTWSNGARCARCAYTTYGGNYTWSNTYNA 713
190 rLeuArgProValProAspAlaLeuLysSerLeuTrpTyrLysAsnLeu 207
714 YTGAAVACARGTNCARGNGCCNCAARCCNMNGNTGCCAYARAAYTYNA 763
207 hrcGlyProGluAsnLeuThrLeuAsnHisThrAspLeuValProCysLeu 223
764 CNGCNCNCARATATATACNTYNAAYCAYACNGAYTYTNGTNCNTGYTN 813
224 CysLleGlnValTrpSerLeuGluProAspSerGluArgValGluPheCys 240
814 TGTATTCARGTGTGCCNYTNGARCCNGAYWSNGTNGMACNAAYATHTG 863
240 sProPheArgGluAspProGlyAlaHisArgAsnLeuTrpHisLleAla 257
864 YCCGTYTWSNGARARAYCCNMGNCACARARAYTYTNGCARCGCNCNM 913
257 rGluArgValLeuSerProGlyValTrpGlnLeuAspAlaProCysCys 273
914 GNTNMGNTYNTNACNTNCARWSNTGTYTNGAYCGCNCNTGYWSN 963
274 LeuProGlyLysValThrLeuCysTrpGlnAlaProAspGlnSerProCys 290
964 YTNCCNCNCARGCNCNTYNTGYTGGMNCNCNCNGCNGCNGAYCCNTG 1013
290 sGlnProLeuValProProValProGlnLysAsnAlaThrValAsnGluPro 307
1014 YCARCCYTYTNGTNCNCNTYNTWSNTGGARARAGTNCNGTNGAY.... 1059
307 rGlnAspPheGlnLeuValAlaGlnHisProAsnLeuCysValGlnVal 1323
1060 .....Gln 1062
324 SerThrTrpGluLysValGlnLeuGlnAlaCysSerTrpAlaAspSerLe 340
1063 AAATWSMSNGARARAYTNCARTCARCARGTYYTNGCGCNGAYWSNT 1112
340 uGlyProPheLysAspAspMetLeuLeuValGluMetLysThrGlyLeu 357
1113 NGCNCNYTNAARGAYGATGTYTNTYNTGARAACNMGNCNCNCARG 1162
357 sAsnThrSerValCysAlaLeuGluProSerGlyCysThrProLeuPro 373
1163 AYAAAYGMSNTYNTGYCNYTNGARCCNMSNGNTGYACMSNTYNTNCN 1212
374 SerMetLaserThrArgAlaAlaArgLeuGlyGluGluLeuGlnAs 390
1213 WSNARCCNWSNACNMGNCNMGNYTNGNGARTRAYTYNTYTCARGA 1262
390 PheArgSerHisGlnCysMetGlnLeuTrpAspAspAspAsnMetGlys 407
1263 YTYNCAWMSNGARTRGYTNCARITYNTG...GATGAYGAYTYTNGNG 1309
407 erLeuTrpAlaCysProMetAspLysTyrLleHisArg 419
1310 CNYTNGCGCNTGYCCNATGAYARAATATHTCAAYAR 1347
Seq_name: /SIDS1/9cgcdata/geneseq/geneseqn-emb1/NA2001A.DAT: AAC85030
Seq_documentation_block:
ID AAC85030 standard; cDNA; 1725 BP.
XX
XX AAC85030;
XX
XX 08-MAY-2001 (first entry)
XX

```

```

DE Human zcytor14 variant zcytor14-1 degenerate nucleotide sequence.
XX
XX Cytokine receptor; zcytor14; human; inflammation; rheumatoid arthritis;
KW antiinflammatory; gene therapy; vaccine; variant; zcytor14-1; ss.
XX
XX Homo sapiens.
XX
XX NO200104304-A1.
XX
XX 18-JAN-2001.
XX
XX 30-JUN-2000; 2000MO-US18383.
XX
XX 07-JUL-1999; 9905-0348854.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX
XX Presnell SR, Burkhead SK, Powder SL;
XX
XX WPI; 2001-112618/12.
XX
XX P-PSDB; AAB61881.
XX
XX New polypeptide encoding a human cytokine receptor zcytor14, for
XX treating inflammation e.g. rheumatoid arthritis -
XX
XX Disclosure; Page 101; 112pp; English.
XX
XX The invention provides a new human cytokine receptor designated zcytor14.
XX zcytor14 can be expressed by standard recombinant methodology. The
XX encoding nucleic acid is useful for detecting the expression of a
XX zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be
XX used to screen biological samples in vitro for the presence of zcytor14.
XX proteins, polypeptides and peptides having zcytor14 activity can be
XX administered to a subject who lacks an adequate amount of this
XX polypeptide, for treating inflammation and conditions such as rheumatoid
XX arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14
XX antibodies) can be used to treat a subject who produces an excess of
XX zcytor14. zcytor14 nucleotide sequences can also be used to provide
XX zcytor14 to a subject. The present sequence represents a degenerate
XX nucleotide sequence of a variant of the human cytokine receptor zcytor14,
XX designated zcytor14-1. The variant is a truncated form of the receptor
XX polypeptide and lacks amino acid residues 1-113 of zcytor14.
XX
XX Sequence 1725 BP; 192 A; 226 C; 332 G; 216 T; 759 other;
XX
XX
XX alignment_scores:
XX      Quality: 852.50      Length: 339
XX      Ratio: 3.947      Gaps: 3
XX      Percent Similarity: 63.717      Percent Identity: 48.968
XX
XX alignment_block:
XX US-09-899-471-2_COPY_21_440 x AAC85030
XX
XX Align seg 1/1 to: AAC85030 from: 1 to: 1725.
XX
XX 90 GlnGluSerArgAsnAlaSerLeuGlnAlaGlnValValLeuSerPheG 106
XX ::::||||| |||||::: ||||| |||||::: |||||
XX 1 GARGARCCNMGNAAYGCMWSNTNCARGCNCARGTNGTYTWSNTTYCA 50
XX ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 106 nLatYrProLleAlaArgCysAlaLeuLeuGlnValGlnValProAla 123
XX ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 51 RGCNTAYCCNACNCGNMGNTGYTYTNTYNGARGTNCARGTNCNGCNG 100
XX ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 123 splEuValGlnProGlyGlnSerValGlySerAlaValAlaPheAspCys 139
XX ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 101 CNYTNGTNCARTTYGGNCARWSNGTWSNGTNGTNTAYGAYTYGTYTY 150
XX ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 140 GluAlaSerLeuGlyAlaGluValGlnLleTrpSerTyrThrLysProAr 156
XX ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 151 GARGCNCNTYNGCNGMSNGARGTNGNATHTGGWSNTATAYACARCCNMG 200
XX ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 156 gTYrGlnLysGluLeuAsnLeuThrGlnGlnLeuPro..... 168

```


1 LeuGluValArgLeuMetGluIupProGlnAspThrAlaIArgCysSerLeuGlyLeu 17
 259 CTGGAGAGGCGTGTGGGGCCCTCGAGAGCGCTAACCCACCTCTCTCCGGGCGCT 308
 17 userCysHisLeuTrpAspPheLysAspValLeuCysLeuProGlySerLeuG 34
 309 CTCCTGGCCGCTCTGGGAGCAGTGCACATACCTGTGGCTGGCTGGGGACATG 358
 34 InsertAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeu 50
 359 TCCGTGCTCCGGGCCCCCGGTGCTGGCGGCTTACGCACCTCCAGCACAGAGCTG 408
 51 ValLeuArgCysProGlnIupTrpAspCysAlaLeuArgValArgValVa 67
 409 GTGGCGAGGTGCCAGAGAGAGACCGACTGTGCATCTGTCTGTGGTGGGC 458
 67 ValHisLeuAlaValHisGlyHisTrpAlaGluProGluGlu..... 81
 459 TGTCCACTGTGGCGGCTGCATATGGGCACCTGGAGAGAGCCCTGAACATGAGAA 508
 82AlaGlyLysSerAspSerGluLeuGlnGluSerArgAspAlaSer 96
 509 AGTTTGGAGGACAGCTGCACTAGGGGTGGAGGAGCCATGAGMATCCCTCT 558
 97 LeuGlnAlaGlnValValLeuSerPheGlnAlaArgProIleAlaArgC 113
 559 CTCGAGGCCCAAGTGGTGGCTCTCTCTCCAGGCCATACCTACTGCCCGCTG 608
 113 sAlaLeuLeuGluValGlnValProAlaAspLeuValGlnProGlyGlns 130
 609 CGCTCTGTGGGAGGTCAGAGTGGCTCTGCTGCTTGTGCAGATTGGTGCAGT 658
 130 erValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGlu 146
 659 CTGGGGCTCTGTGGATATAGACTGCTTCAGAGCTGCCCTTAGGGAGTAG 708
 147 ValGlnIleTrpSerTrpThrLysProArgTyrGlnLysGluLeuAsnLe 163
 709 GTAGAAATCTGGTCTCATCTACTAGCCAGGTACGAGAAAGAACTCAACCA 758
 163 uThrGlnGlnLeuPro..... 168
 759 CACACAGCAGCTGCCTGACTGACAGGGGCTCGAAGTCTGGAACAGCATCC 808
 169 808
 809 CGAGCTGCTGGGCCCTGGCCCTGGCTACAGTGCACGACATGGTGCACAC 858
 173 ValLeuLeuThrLeuAspValSerGluGlnGlnAspPheSerPheLeu 189
 859 GTGCATCTGGTTCTAAGTGTCTCTGAGGAGCACACTCTGGCCCTCTCCCT 908
 189 uTrpLeuArgProValProAspAlaLeuLysSerLeuTrpTyrIleAsnL 206
 909 GTACTGGAATCAGGTCCAGAGGCCGCCCAAAACCCCGGTGGCACAAACCC 958
 206 euThrGlyProGlnAsnIleThrLeuAsnHisThrAspLeuValPro.Cy 222
 959 TGTGTAGGAGCTCCCCCTTCCCAAGTC...CATTCACACTGTAGGCGCATG 1005
 222 sLeuCys.....IleGlnValTrpSerLeuGlnProAsps 234
 1006 CCGTGTGCAAGAGCAGCAGTGCATATCAGAGAGATCCCTTGAAAGAGACT 1055
 234 erGluValGluIupPheCysProPheArgGluAspProGluAlaHisArg 250

1056	CA	CCCCACACGAGGAA	1072
251	AsnLeu	TrpHisIleAlaArgLeu	ArgValLeuSerProGlyValIrrpG1	267
1073	AATTCG	TTG		1081
267	nLeu	AspAlaProCysCysLeuProGly		276
1082	GGGACCTTCCTCCCTCGGT		1102
seq_name: /SIDSI/gcgdata/geneseq/geneseqn_emb1/NA2001B.DAT: AAS78822				
seq_documentation_block:				
XX	ID	AAS78822	standard; cDNA; 980 BP.	
XX	AC	AAS78822;		
XX	DT	13-FEB-2002	(first entry)	
XX	DE	DNA encoding novel human diagnostic protein #14626.		
KW	Human; Chromosome mapping; gene mapping; gene therapy; forensic;			
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.			
XX	OS	Homo sapiens.		
XX	PN	WO200175067-A2.		
XX	PD	11-OCT-2001.		
XX	PF	30-MAR-2001; 2001MO-US08631.		
XX	PR	31-MAR-2000; 2000US-0540217.		
XX	PR	23-AUG-2000; 2000US-0649167.		
XX	PA	(HYSE-) HYSEQ INC.		
XX	PI	Dmanac RT, Liu C, Tang YT;		
DR	WPI: 2001-639362/73.			
DR	P-FSDB: ABG14635.			
XX	New isolated polynucleotide and encoded polypeptides, useful in			
PT	diagnostics, forensics, gene mapping, identification of mutations			
PT	responsible for genetic disorders or other traits and to assess			
PT	biodiversity			
XX	Claim 1; SEQ ID NO 14626; 103pp: English.			
XX	The invention relates to isolated polynucleotide (I) and			
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,			
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome			
CC	and gene mapping, and in recombinant production of (II). The			
CC	polynucleotides are also used in diagnostics as expressed sequence tags			
CC	for identifying expressed genes. (I) is useful in gene therapy techniques			
CC	to restore normal activity of (II) or to treat disease states involving			
CC	(II). (II) is useful for generating antipodest against it, detecting or			
CC	quantitating a polypeptide in tissue, as molecular weight markers and as			
CC	a food supplement. (II) and its binding partners are useful in medical			
CC	imaging of sites expressing (II). (I) and (II) are useful for treating			
CC	disorders involving aberrant protein expression or biological activity.			
CC	The polypeptide and polynucleotide sequences have applications in			
CC	diagnostics, forensics, gene mapping, identification of mutations			
CC	responsible for genetic disorders or other traits to assess biodiversity			
CC	and to produce other types of data and products dependent on DNA and			
CC	amino acid sequences. AAS64197-AAS94564 represent novel human			
CC	diagnostic coding sequences of the invention.			
CC	Note: The sequence data for this patent did not appear in the printed			
CC	specification, but was obtained in electronic format directly from WIPO			
CC	at ftp://ipo.int/pub/published_pct_sequences.			
XX	Sequence 980 BP; 182 A; 313 G; 315 G; 170 T; 0 other;			

Alignment_scores:

Quality: 596.00 Length: 175
Ratio: 4.000 Gaps: 4
Percent Similarity: 85.143 Percent Identity: 69.714

Alignment_block:

US-09-899-471-2_COPY_21_440 x AAS78822/rev ..

Align seg 1/1 to reverse of: AAS78822 from: 1 to: 980

```
1 LeuGluArgLeuMetGluProGlnAspThrAlaArg.CysSerLeuGluYL 17
|||||..... |||||||..... |||||
720 CTGGAGAGGCTGTGGGCTTCAGACCGCTACCCACTGCTCCGGGCC 671
17 euserCysHisLeuTrpAspGlyAspValLeuCysLeuProGlySerLeu 33
|||||..... |||||||..... |||||||.....
670 TCTCTGCCGCTGGGACAGTACATCTGCTGCTGGGACATC 621
34 GlnSerAlaProGlyProValLeuVal.Pro.ThrArgLeuGlnThrGlu 49
|||||..... |||||||..... |||||||
620 GTGCTGCTGGGCGCCGCGCTGTGGCTTACCACTTGAGACAGAG 571
50 LeuValLeuArgCysProGlnLysThrAspCysAlaLeuArgValArgVa 66
|||||..... |||||||..... |||||||
570 CTGGTGTGAGAGTCCAGAGAGACCGACTGTGACCTGTCTGCGGT 521
66 lValValHisLeuAlaValHisGlyHisTrpAlaGluProGluGlu... 81
|||||..... |||||||..... |||||||
520 GGGTGTCCACTTGGCGCTGCATGGGCACTGGGAGAGCCTGAAGATGAG 471
82 .....AlaGlyLysSerAspSerGluLeuGlnLysSerArgAsnAla 95
|||||..... |||||||..... |||||||
470 AAAAGTTGGAGAGACACTGACTTAGGGGTGAGAGACCTAGAAATGCC 421
96 SerLeuGlnAlaGlnValValLeuSerPheGlnAlaTyrProIleAlaAr 112
|||||..... |||||||..... |||||||
420 TCTCTCCAGGCCCAAGTGTCTCTCTCCAGGCCCTACCTACTGCCCG 371
112 gCysAlaLeuLeuGlnValGlnValProAlaAspLeuValGlnProGlyG 129
|||||..... |||||||..... |||||||
370 CTGGCTCTGCTGGAGGTGCAAGTCTGCTGCCCTTGTGCAATTGGTC 321
129 lnservAlaGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAla 145
|||||..... |||||||..... |||||||
320 AGTCTGTGGGCTGTGTGTATATGACTGCTTGAGGCTGCCCTAGGAGAT 271
146 GluValGlnIleTrpSerTyrThrLysProArgTyrGlnLysGluLeuAs 162
|||||..... |||||||..... |||||||
270 GAGGTAGGAATCTGTCTATATCTACAGCCAGGTACGAGAGAACTCA 221
162 nLeuThrGlnGlnLeuPro 168
| |||||||
220 CCACACAGAGAGCTGCT 202
```



```

330 InLeuGlnAlaCys.....SerTIPalaasp 338
    |||
995 GCGAGTGAATGTCCCGACAGAGTGGCTCTCCCATCTGGAGCTGTG 1044
339 SerLeuGlyProPheLysAspAspMetLeuLeuValGlnMetLysThrG1 355
    |||
1045 AGCATGGATATCC...CAGGCCGACGAGCTGACGCTTTCCTTCTGAG 1091
355 YLeuAsnThrSerValCysAlaLeuGlnProSerGlyCys..... 369
    |||
1092 GACATATGCCACCTTCAGTGTGCTGTGAGTGAACCCAGGTTTGGGGCCG 1141
370 ..ThProLeuProSerMetAlaSerThrArgAlaAlaArgLeuGlyGlu 385
    |||
1142 ATACCCCATGCTCTCTGTGTACAGATCAAGCCAGACCCAGGCTCAATC 1191
386 GlnLeuLeuGlnAsp.....PheArgSerHisGlnCys 397
    |||
1192 CCAGTGAGCGCTAGACCTCATTCATCCCTTCCCTGAGGACAGATTCAT 1241
397 tGlnLeuTrpAsnAspAspAspMetLysSerLeuTrp.....AlaC 411
    |||
1242 CCTGTGTGGAGGTGAGAT...GTCCATTTTGCTGAGACACGCTTGT 1288
411 YSPrometAsp 414
    |||
1289 GTCCGTGATGAC 1299

```

```

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-09-188-930-6

```

```

seq_documentation_block:
: Sequence 6, Application US/09188930A
: Patent No. 6150502

```

```

GENERAL INFORMATION:
: APPLICANT: Watson, James D.
: APPLICANT: Strachan, Lorna
: APPLICANT: Sleeman, Matthew
: APPLICANT: Murist, Rene
: APPLICANT: Murlson, James Greg
: TITLE OF INVENTION: Compositions Isolated From Skin Cells
: FILE REFERENCE: 11000.1011C1
: CURRENT APPLICATION NUMBER: US/09/188,930A
: CURRENT FILING DATE: 1998-11-09
: NUMBER OF SEQ ID NOS: 348
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 6
: LENGTH: 1059
: TYPE: DNA
: ORGANISM: Mouse
US-09-188-930-6

```

```

alignment_scores:
Quality: 149.00      Length: 152
Ratio: 1.795         Gaps: 6
Percent Similarity: 54.605      Percent Identity: 28.289

```

```

alignment_block:
US-09-899-471-2_COPY_21_440 x US-09-188-930-6

```

```

Align seg 1/1 to: US-09-188-930-6 from: 1 to: 1059

```

```

204 LysAsnLeuThrGlyProGlnAsnLeuThrLeuAsnHisThrAspLeuVal 220
    |||
616 AAATTTGTCTGTGAGGCCACACTGTGAGACTGCTTATGTAATTCCTCT 665
220 LProCysLeuGlyLeuGlnValTrpSerLeuGlnProAspSerGlyArgV 237
    |||
666 GCCCGCATGTGCATAGAGCCCTCTACTGTGCAAGAGACACTGTGAGGC 715
237 aGlnPheCysProPheArgGlnAspProGlyAlaHis...ArgAsnLeu 252
    |||

```

```

716 GCAAAAGTGTCCCTTCCAGAGTGGCTGTGAAGCTTATGGCTCAAGCTTC 765
253 TrpHisIleAlaArgLeuArgValLeuSerProGly.....ValTr 266
    |||
766 TGGCAGTCAATATGCTTCACTGACTGACTACAGACAGACATCATGATGTCT 815
266 pGlnLeuAspAlaProCysCysLeuProGlyLysValThrLeuCysTrpG 283
    |||
816 GGCTCTGACACTCCGCTGCCACCTGAAACTGAGAGCCCTCTCTCTGGA 865
283 InAlaProAspGlnSerProCysGlnProLeuValProProValProGln 299
    |||
866 GGCAGACCCACTCAACCCCTGCGAACCCTT..... 897
300 LysAsnAlaThrValAsnGlnProGlnAspPheGlnLeuValAlaGly.. 315
    |||
898 CCACAGCCACAGACAGAGAGTCAAGAGATATATCTCGAATAAT 947
316 .....HisProAsnLeuGlyValGlnValSerThrTrpLysValG 330
    |||
948 GGACTTGACACCCCGACCTGTGTTAAGTCTCA...TTGAAACACAGA 994
330 InLeuGlnAlaCys.....SerTIPalaasp 338
    |||
995 GCCAGCTTGATGTCCCGACAGAGTGGCTCTGCCATTCCTGGAGCTGTG 1044
1045 AGCATG 1050

```

```

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-342-930-1

```

```

seq_documentation_block:
: Sequence 1, Application US/08342930
: Patent No. 5821084

```

```

GENERAL INFORMATION:
: APPLICANT: OLAMSTED, ELIZABETH A.
: APPLICANT: MAURO, LAURA J.
: APPLICANT: DAVIS, ALAN R.
: APPLICANT: DIXON, JACK E.
: TITLE OF INVENTION: OSTEOBLAST-TESTICULAR PROTEIN TYROSINE
: TITLE OF INVENTION: PHOSPHATASE
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESS: MORRISON & FOSTER
: STREET: 755 Page Mill Road
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304-1018
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/342,930
: FILING DATE: 21-NOV-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: KONSKI, ANTOINETTE F.
: REGISTRATION NUMBER: 34,202
: REFERENCE/DOCKET NUMBER: 20344-20975.00
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 813-5600
: TELEFAX: (415) 494-0792
: TELEX: 706141
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5455 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear

```


4908 CATTGTGACGTCGTTAAGGAAGCAATGACACCCAGTTAGCCAA 4957

1957

107 AlaTyrProIleAlaArgCysAlaLeuLeuGluValGlnValProIleAs 123

3995	GCTTATCAATAATGAGAGAGACACCTTTCTGGACCTCAGTTG...	4035
123	pleuValGlnProGlyGlnSerValGlySerAlaValAlaPheaspCysPheG	140
4036TTGTAATATGCTGTGTTT..	4053
140	IuAlaSerLeuGlyAlaGluValGlnIleTyrSerTyrThrLysProArg	156
4054ATATGGCTCTTAT	4065
157	TyrGlnLysGluLeuAsnLeuThrGlnGlnLeuPro.....	168
4066CTTAACCAACTGGTACTTCCCTTATATGCTGGGCTT	4103
169	.AspGlyAspAsnValLeuLeuThrLeuaspValSergLugluGlnAsp	185
4104	CTTGGCCCGCTTACTCTTCCCTCCCAAAAAGTGTGGC.....	4140
185	heserPheLeuLeuTyrLeuArgProValProaspAlaLeuLysSerLeu	201
4141GCCCCGTGGTCACTGGGCTT.....CTA	4164
202	TrpTyrLysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAs	218
4165	TGTTGGTGGCTCCCTCTTCAACCGCTTCCAGCCG.....	4200
218	pleuValProCysLeuCysIle.GlnValTyrPheSerLeuGlnProaspSer	234
4201TTGTCTCTGTTCCACTGCTGTCTGGTCAACCGTCAACGG	4237
235	GluArgValGluPheCysProPheArgGluaspProGlyAlaHisArgAs	251
4238	GAAACGCGATATCTTGTGGTTCTTACATCATCAGCCG.....	4273
251	pleuThrHisIleAlaArgLeuArgValLeuSerProGlyValTyrPheL	268
4274CTGGACGCTTCTTACTTCTTGAGATGTGGTTGC..GAGG	4310
268	euaSpAlaProCysCysLeu.....Pro	275
4311	CTATCCACCATGTTGTTGTGGCCAGGTTCCTAGTGTACCAAGTCCCG	4360
276	Gly.....LysValThrLeuCysThrPheL	283
4361	GCTGGGCTATTGGCACAGTACTAGCGGATGACAGCATCCACATGCTGGCT	4410
283	nalAlaProaspGlnSerProCysGlnProLeuValProProValProGlnL	300
4411	GCTGGCCCTCGGTCAACCCCTGTACTGTGATGTGTTCCG..CCACAGG	4457
300	ysaSnAlaThr.....	303
4458	TCGCTTTGAACAGACTTCATGATCAATACTCTCGAGGAGAGATGA	4507
304ValasnGluProGlnAspPheGlnLeuValAlaGlyHis.....	316
4508	AAAGAGTGTACCGGCTCACTCAACCCGCGTTAAAGACAAACCAATTACC	4557
317ProasnLeu.....	319
4558	CAGGAGAGTCACTCAACATTAAGTACCTCCCTCACTGATGATGACTTCA	4607
320CysVal.....	321
4608	ATTCTCTCTGATGTGGTGACTGTGGGCGGTCCGATCGGCATGAATC	4657
322GlnAlaSerThrTyrLysVal.....	329
4658	TCGGTCCCGCTTCACAAAGTTTTCAAATGTGGGCGAGTATCGTAACCTCTT	4707
329	329


```

4708 AATGATCCTTGCAAGTCGATCGTGCCTGCTGCTAGTCCAGACTAAT 4757
330 .....GlnLeuGlnAla 333
4758 GCCAAACTGCTGATTTTTCGGTTGANAAGATGACAGCTGGAGACC 4807
334 CysSerTrp..AlaAspSerLeu.....GlyProPhelysAspasp 346
4808 GTGTGTGTGTATGACGGTCTGACCGCATGCTCACTTCAAAAGACAT 4857
347 MetLeuLeuValGlnMetLysThrGlyLeuAsnAsnThrSerValCysAl 363
4858 TTGGTGCTGCTTCTTGGACCAACCAAGTAGTAGCGGTTTAGGTGCAC 4907
363 Aleu.....GluProSerGlyCysThrProLeuProSerM 375
4908 CATTTGTGACGTCGTTAAGAGAGACCAATGACCCCAAGTTAGCCAA 4957
375 eAlaSerThrArgAlaAlaArg.....LeuGlyGlnGlnLeuLeu 388
4958 TGCCACAGCAGGAGAGCCGCCAGCGCTGCTAAAGGTCTCAGTTGGAG 5007
389 GlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsnAspasp 403
5008 TGGGACCTTCACCAAGAGAGAGAGAGAGAGCGCGTATGATGAT 5052

Seq.name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:us-09-017-631-23

Seq.documentation_block:
Sequence 23, Application US/09017631B
Patent No. 6153407
GENERAL INFORMATION:
APPLICANT: Sytkowski, Arthur J.
APPLICANT: Grodberg, Jennifer
TITLE OF INVENTION: ERYTHROPOIETIN WITH ALTERED BIOLOGICAL
FILE REFERENCE: NEDH92-04A22A
CURRENT APPLICATION NUMBER: US/09/017,631B
CURRENT FILING DATE: 1998-02-03
EARLIER APPLICATION NUMBER: US 08/808,881
EARLIER FILING DATE: 1997-02-28
EARLIER APPLICATION NUMBER: US 08/383,743
EARLIER FILING DATE: 1995-02-02
EARLIER FILING DATE: 1993-08-26
EARLIER APPLICATION NUMBER: US 07/920,810
EARLIER FILING DATE: 1992-07-28
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 3601
TYPE: DNA
ORGANISM: Homo sapien
US-09-017-631-23

Alignment_scores:
Quality: 101.50 Length: 504
Ratio: 0.495 Gaps: 34
Percent Similarity: 40.675 Percent Identity: 22.024

Alignment_block:
US-09-899-471-2_copy_21_440 x US-09-017-631-23/rev ..
Align seg 1/1 to reverse of: US-09-017-631-23 from: 1 to: 3601

13 CysSerLeuGlyLeuSerCys.....HisLeuTrpAspGlyAspVa 26
|||||
3015 TGCCTCTGAAATGCTTCCTGCTGCGCCCTCAAGTTGGCCCTGTGACAT 2966
26 Ileu.....CysLeuProGlySerLeuGlnSerAlaProGlyP 39
|||||
2965 CCTTGATCTCAGAGTTGC.....TCTTGAGACAGTTCTCTGACC 2925

```

```

39 ro.....ValLeuValProThrArgLeu.....GlnThrGlnLeu 50
|||
2924 CCGTAGATGTCATTTGCTGCGACTGTGAGTGTCCATGGAGACAGCGTGGCGT 2875
51 ValLeuArgCysProGlnLys.....ThrAs 59
|||||
2874 GAGCTGAGAGCCCTCGACGGGGTTCCAGAGAGTGGGGGAGAGGTGTGGC 2825
59 pCysAlaLeuArgValArgValValHisLeuAlaValHisGlyHis 75
|||||
2824 ACAAGCAATGTTGGTAGAGGAGGTGTGTGATATGCCACGTGACACACC 2775
76 TrpAlaGlu...ProGlnGluAlaGlyLysSerAspSerGlnLeuGlnG 91
|||||
2774 TGGTCATCTGTCCTCCCTGCTCCGACGC..... 2748
91 uSerArgAsnAlaSerLeuGlnAlaGlnValValLeuSerPheGlnAla 108
2747 .....C 2747
108 YrPrOleAlaArgCysAlaLeuLeuGluValGlnValProAlaAspLeu 124
|||||
2746 TCCCTCTGTACACCTTCACCTTCCCGGAGAAATGGAGTACAGCTCG 2697
125 ValGlnProGlyGlnSerValGlySerAlaValPheAspCysPheGlu.. 140
|||
2696 GAGAGATTGGCGGAAGTGTACAGAGT.....GATTTTCGAGAGT 2656
141 AlaSerLeuGlyAlaGluValGlnIleTrp.....SerTyrThrL 154
|||||
2655 GAGCAGCTGAGGC.....CGCATCTGAGGGGAGCATGCTTCTTCTTC 2615
154 YsPrArgTyrGlnLysGlnLeuAsnLeuThrGlnGlnLeuProAspGly 170
|||||
2614 TGCCAGAGAGAAACAGAGAGGTCGTCAGTCCACAG..... 2577
171 AspAsnValLeuLeuThrLeuAspValSerGlnGlnAspPheSerPh 187
2576 .....AAAGGAAGATATACGAGCA 2557
187 eLeuLeuTyr.....LeuArgProValProAspAlaLeuLysSerLeu 202
|||||
2556 GTTCCCTCTACTCCTTAGCAGACACCTTCCCTTCTTACAGAAAG.... 2511
202 rPtyrLysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAsp 218
2510 .....GGCAGACGAGAG..... 2499
219 LeuValProCysLeuCysIleGlnValTrpSerLeuGlnProAspSerG 235
|||||
2498 .....TGTCCGCTCTACTACTACTGCTGCTCCACAGCCCGAAGCAG 2458
235 uArgValGluPhe.....CysProPhea 243
|||||
2457 AGTGTGAGGCTGCCAAGGCCACGACGCGTTATTCACATCAGCTGCA 2408
243 rGlnAspProGlyAlaHis.....ArgAsnLeuTrp 253
|||||
2407 GG.....GGCTCCACGCGTGGAGAGACTTACCAACAGGGCGCTGG 2367
254 HisIleAlaArgLeuArgValValLeuSerProGlyValTrpGlnLeuAsp 270
|||||
2366 CCGCGCAGAGACGCTTCGAGACAGAGGCCACGCGCTGCCAATCTTCTTC 2317
270 lArProCysLysLeuProGlyLysValThrLeuCysTrpGlnAlaProAsp 286
|||||
2316 GGCTCTGTCGCCGCTACACAGAGAGTGCATC...TGGAGTACAGCTGAC 2270
287 GlnSerProCysGlnProLeuValProProValProGlnLysAsnAlaHis 303
|||||
2269 .....CCATGTACCCCTCTC...CTCTGCGCCCTCAGACAGCCCAAGC 2229

```



```

1      REGISTRATION NUMBER: 40,261
2      REFERENCE/DOCKET NUMBER: 78441-062
3      TELECOMMUNICATION INFORMATION:
4      TELEPHONE: (416) 364-7311
5      TELEFAX: (416) 361-1398
6      INFORMATION FOR SEQ ID NO: 33:
7      SEQUENCE CHARACTERISTICS:
8      LENGTH: 3602 base pairs
9      TYPE: nucleic acid
10     STRANDEDNESS: single
11     TOPOLOGY: linear
12     MOLECULE TYPE: DNA (genomic)
13     ORIGINAL SOURCE:
14     ORGANISM: Homo sapiens
15     FEATURE:
16     NAME/KEY: CDS
17     LOCATION: join(625..637, 1201..1346, 1605..1691, 2303..2482)
18     LOCATION: 2617..2772)
19     FEATURE:
20     NAME/KEY: mRNA
21     LOCATION: join(625..637, 1201..1346, 1605..1691, 2303..2482)
22     LOCATION: 2617..2772)
23
24     US-08-883-795A-33

```

```
alignment_scores:
  quality: 101.50      Length: 504
  Ratio: 0.495         Gaps: 34
Percent Similarity: 40.675  Percent Identity: 22.024
```

```
alignment_block:
```

```

US-09-899-4/1-2_COPY_21_440 x US-08-883-795A-33/rev ..
Align seg 1/1 to reverse of: US-08-883-795A-33 from: 1 to: 3602

```

```

13 CysSerLeuGlyLeuSerCys.....HisLeuTripsArgLysAlaPva 26
   |||||.....|||||
3016 TGCTCTCGAATGCCTTCCGTCTGTGGGCCCTCAAGTTGCCCTGTGCAT 2967
   ..||..
26 Ileu.....CysLeuProGlySerLeuGlnSeralAlaProGly 39
   :|||      |||
   |||||.....:|||||
2966 CCTTAGATCATTCAGAGTTC.....TCTCGACAGTGTCTCTGGCC 2926
   .....
39 ro.....ValIeuValProThrArgLeu.....GlnTrgInLeu 50
   |||||.....:|||||
2925 CCTAGATGTAATTCCTGGCACCTGAGATGTCCATTGGACAGCTGGCGCT 2876
   .....
51 ValIeuArgCysProGlnLys.....Thras 59
   |||||.....:|
2875 GAGCTGAGAGCCCCCGACAGGGGTTTCAGAGAAGTGCGGGGAGGCTGTGGC 2826
   .....
59 pCysAlaLeuArgValArgValAlaValHisLeuAlaValHisGlyHis 75
   |||||.....:|||
2825 ACAAGCAATGTGGTGAGGAGAGGTGTGATATGCCAGGTGGACACACC 2776
   .....
76 Triplaclu...ProGluGlnAlaGlyLysSerAspSecGluLeuGlnI 91
   |||:.....|||
2775 TGTGCATCTGCCCCCTGTCTCCAGGC..... 2749
   .....
91 uSerArgAsnAlaSerLeuGlnAlaGlnValValLeuSerPheGlnAlar 108
   .....
2748 .....C 2748
   .....
108 yrProIleAlaArgCysAlaLeuLeuGlnValGlnValProAlaaspLeu 124
   .....
2747 TCCCCTGTGACACTTCAGCTTCCCGGAGAAATTGGAGACATCG 2698
   .....
125 ValGlnProGlyGlnSerValGlySeralValaPheAspCysPheGlu.. 140
   :|||:.....:|||||
2697 GAGAGATTGGCGGAAGTGTACACGT.....GATTGTTCGGAGTG 2657
   .....
141 AlaSerLeuGlyAlaGluValGlnLeuTrp.....SerTryThrL 154
   :|||:.....:|||||

```


2656	GAGCAGCTGAGGC.....CGCATCTGGAGGGAGATGGCTTCCCTC	2616
154	ysProrArgTyrGlnIysGlnLeuAsnLeuThrGlnGlnLeuProaspGly	170
2615	TGCCAAGAGAAACAAAGAGGCGTCTGACAGTGCACAG.....	2578
171	AspAsnValLeuLeuThrLeuAspValSerGluGluInAspPheSerP	187
2577AAAGGAGGAAATACGGACA	2558
187	eLeuLeuTyr.....LeuArgProValProAspAlaLeuIysSerLeuT	202
2557	GTTCCTGTACTCTTACGAAAGACCTTCTCCCTCTTACAGAAAG...	2512
202	rPtyrIysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAsp	218
2511GGCAACCAAG.....	2500
219	LeuValProCysLeuCysIleGlnValTyrSerLeuGluProAspSerG	235
2499TTCGGCTCTCTACTCACTGGGCTCCAGACCCGAGACAG	2459
235	uArgValGluPhe.....CysProPheA	243
2458	AGTGTGTAGGTGGCCAAAGGCCACTGACGGCTTTATCCATGCAGCTGCA	2409
243	rgGluAspProGlyAlaHis.....ArgAsnLeuTyr	253
2408	GG.....GGCTCCCAAGCGCTGGGAAGAGTTGACCAACAGGCGCTCG	2368
254	HisIleAlaArgLeuArgValLeuSerProGlyValTyrGlnLeuAspA	270
2367	CCCGGAGAGACAGCTTCGCACACAGCAGGCGCCCTGCCAGACTTCTAC	2318
270	laprCysCysLeuProGlyIysValThrLeuCysTyrGlnAlaProAsp	286
2317	GGCCTGTGCGCCGACCTACAGGAGTGAGTCTC...TGGAGTGCAGTGCAG	2271
287	GlnSerProCysGlnProLeuValProProValProGlnIysAsnAlaThr	303
2270CCATCTACCCCTTC...CCTCTGCCCCACAGAGCCCCAACG	2230
303	r.....ValAsnGluProGlnAspPheGlnLeuVal.....	313
2229	ACCAAGCTGACCAACACAGAGTATGCAATTAAGACTTGTGAATGAATGA	2180
314AlaGlyHis	316
2179	ATGATGATGATGATGATGATGATGATGATGATGATGATGATTCAT	2130
317	ProAsnLeuCysValGlnValSerThr.....	325
2129	ACAGCCCTCATATTATTCTTTCTTTCTTTCTTTTGGACAGAGGGCTC	2080
326TyrGluIysVal.....GlnLeuGlnAlaC	334
2079	ACTGTTCACATGAGGCTTGAGTGCAGTGGTGTGATGCACAGCTACCTGACG	2030
334	ys.....SerTyrAlaAspSerLeuGlyProPheIysAspAspMetLeu	348
2029	CCTCAAATTCCTGGGCTCAAGAGATCTCTCCGCCCTCAGCCTTCCAATAT	1980
349	Leu.....	349
1979	CTGGAGCTACCAACCATGCACACACTTACCTGACTAATTTTAAATGTT	1930
350	ValGluMet.....	352
1929	TGTAGAGATGGGAGATCTCACTATGCTGCTAGCTTGATGCTGAACCTCA	1880
353LysThrGlyLeuAsnAsn	358
1879	GGGCTCAAGCAATTTCTCCCATCTGGGCATCTTCAGCCTGGATTATAGAC	1830

359 ThrSerValLysAla.LeuGluProSerGlyCys.....ThyrProL 372
:::||||:::||||| |||
1829 GTAGGCGCTCGGGCCAGGAGCGCTCATCTCGTCTCCATTTTACCTT 1780
372 eudPro...SerMetAlaSerThyrArgAlaAlaArgLeuGlyGluGluLeu 387
::||| |||::: :: |||||:::||||:::
1779 TCCCGCGATATTCTCCCTTTCATCCAAATCAGGCTCGCAATGAGAT 1730
388 LeuGln 389
||| |||
1729 CTCCAA 1724

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-018-138-1

```
seq_documentation_block:
```

```

Sequence 1, Application US/09018138
Patent No. 6187564
GENERAL INFORMATION:
APPLICANT: Sytkowski, Arthur J.
TITLE OF INVENTION: Production and Use of Recombinant
TITLE OF INVENTION: Protein Multimers with Altered Biological Activity
FILE REFERENCE: BI97-05A
CURRENT APPLICATION NUMBER: US/09/018,138
CURRENT FILING DATE: 1998-02-03
EARLIER APPLICATION NUMBER: 08/890,929
EARLIER FILING DATE: 1997-07-10
NUMBER OF SEQ ID NOS: 24
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 3602
TYPE: DNA
ORGANISM: Human
US-09-018-138-1

```

alignment_scores:	
Quality:	101.50
Ratio:	0.495
Percent Similarity:	40.675
	Length:
	Gaps:
	34
	Percent Identity:
	22.024

```
alignment_block:
    no 00 0000 471 3
```

US-09-899-471-2_COPY_21_440 x US-09-018-138-1/rev

Align seg 1/1 to reverse of: US-09-018-138-1 from: 1 to: 3602

```

13 CysSerLeuGlyLeuSerCys.....HisLeuTrpAspGlyAspVala 26
   ||| ||||| ||||| ||| |||
3016 TGCCTCTCGAATGCTTCCCTGTCTGGGCCCTCAAGTTGACCCCTTGACAT 2967
   .....CysLeuProGlySerLeuInserAlaProGlyP 39
   ||| || | |||||:::||| |||
2966 CCTTAGATTCAGAGTTCC.....TCTCTGACACAGTCTCTGTGCC 2926
   ..ValleuValProThrArgLeu.....GlnThrGluLeu 50
   || |||||:::|| ::|||::|
2925 CCTGAGATGTCATTGCTGGCACACTGTGAGTGTCCATTGACACAGCTTGGCGT 2876
   ValLeuArgCysProGlnLys.....Thrs 59
   |||||:::|||||:::
2875 GAGCTTGAGAGCCCCCTCAGCGGGTTCAAGAGTGGCGGGGAGAGGTGTGGT 2826
   pCysAlaLeuArgValaArgValaIValHisLeuAlaValHisGlyHis 75
   |||::| ||||| ||||| ::|||
2825 ACAAGCAATGTTGGTGAAGAGAGGTGGTGAATGCCCAAGTGAACACACC 2776
   TrpAlaLaL..ProGluGlnAlaGlySerSerpSerGluLeuGlnGl 91
   |||::: || | |||||
2775 TGGCACACTGTGCCCTGTCCGACAGC..... 2749
   userArgAsnAlaSerLeuGlnAlaGlnValValLeuSerPheGlnAlat 108
2748 .....C 2748
```


349	Leu.....	349
1979	CTGGGACTACCAACCATTGCACCACTTCACTTAATTTTAAATGTT	1930
350	*Valgumet.....	352
1929	TGTAGSGAATGGGGGATCTCATACTGCTGCCCTAGGTTGCTGAACATCCA	1880
353 LysThrIglyleuAsnAsn	358
1879	GGGCTCAACAATTCCTCCCATCTCGGGCANCTCAGCCTGGGATTTATAGAC	1830
:	:::	:::
359	ThrservalCysAla LeugluProSerIglyCys..... ThrProl	372
	:::	
1829	GTTAGGCTCTGGGCCAGGACGCTCATCTGCTCCATTTTACCTT	1780
372	eubPro... SermetAlaserThrArgAlaAlalarLeugluIguInluIn	387
	:::	
1779	TCCCCCGGATTCATTCCTCTTCATCCAAATATCCAGGCTCCGCAAAATGAGATT	1730
388	Leugln	389
1729	CTCCAA	1724

```

seq_name:cg21_6/plodata/2/lna/bb_comb.seq:us-08-718-5
seq documentation block:
Sequence 5, Application US/08718388
Patent No. 6271362
GENERAL INFORMATION:
APPLICANT: MORIKAWA, MINORU
APPLICANT: HARADA, NAOKI
TITLE OF INVENTION: GENE ENCODING IGG FC REGION-P-
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSER: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718, 388
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR, GERALD M
REGISTRATION NUMBER: 28, 977
REFERENCE/DOCKET NUMBER: 0230-111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-718-388-5
alignment_scores:
Quality: 100.50
Ratio: 0.498
Percent Similarity: 44.690
Length: 452
Gaps: 30
Percent Identity: 22.345

```


alignment_block:
US-09-899-471-2_COPY_21_440 x US-08-718-388-5 .

Align seg 1/1 to: US-08-718-388-5 from: 1 to: 3661

13	CysSerLeuGluGlyLeuSerCysHisLeuTrpAspGluAspValleu.....	27
1832	TGTGTGGAGGGCTGCACATGGCAGCG.....GGTTTGTCTTAATGTC	1875
28CysLeu.ProGlySerLeuGlnSerAlaProGlyProValleu	41
1876	TGACCGCTGTCTTCCCTCCACACAGCGCTGCAGCTGTGGCCCAATGGCA	1925
42	ValProThrArgLeuGlnThrCldLeuValLeuArgCysProGlnIstb	58
1926	CCATCCACAGAGCGGGCAGTGAAGTTTGGGCGCTGATGGCACCTCTCCAG	1975
58	rAspCysAlaLeuArgValArgVal.....ValHisLeuIav	72
1976	TGGGTGCGCTCGGGCTGAGGGGTGGCTGCTGTGCACACTGGCAG	2025
72	alHisGlyHisTrpAlaIu.....ProGlnIu	81
2026	CTGTGGCG...TGGGTGAAGTGTGTGGCCTCTGCCATCCGGCCAGCG	2072
82	AlaGlyIysSerAspSerGluLeuGlnIuserArgAsnIasSerLeuG	98
2073	GCTGCCAGCCCGTACGACAGCTGAGTCCAGCGTGGGGTACCCCCAT	2122
98	nAlaGlnValValLeuSerPheGlnAla.TyrProIleAlaArgCysAla	114
2123	TACGTCACTGGATGGAGGCACCATTCATTTCCA.....	2157
115	LeuLeuGlnValGlnValProAlaspleuValGlnPro.GlyGlnServ	131
2158	..AGGACCTGGCAGATCTCTGTAGTCCACCTCCAGCGACCAACCT	2204
131	alGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGlnVal	147
2205	TGGGGCGCTGAGAACTTCACTGCTGATGCCATATAGCACCGGGGAGC	2254
148	GlnIleTrpSerTyrThrLys.....ProArgTyrGlnIysG	160
2255	CAGCTGTACACTACACCGCGAGTGCACCTGCACAAATGTACAAACAG	2304
160	uLeuAsnLeuThrGlnGlnLeuPro.....AspGlyAspAsnVal	174
2305	CTTACACACTGAGTGGCCCGCTGGCCCGGAAGTACAGTGGAGCGGCTGT	2354
174	euleuThrLeuAspValSerGluGlnAspPheSerPheLeuTyr	190
2355	TCTGTCACTCTG.....	2365
191	LeuArgProValProAspAlaLeuIysSerLeuTyrTylAsnLeuTh	207
2366CCCTTCACGTGAGCTGCTCTGTCACAGCACACCTGAG	2403
207	rGlyProGlnAsnIleThrLeuAsnHisThrAspleuValProCysLeu	224
2404	CGGGCGCGACCTGGGGTGAGCCACCAACTCCAGGGCTC.....	2440
224	ysIleGlnValTrpSerLeuGlnProAspSerGlu.....	235
2441TGGTGGCTTTCAGCGGGAGACAGCTTCGTGGCGCTG	2476
236	ArgVal.....GluPheCys.....	242
2477	CGCGTGGCGGGCGGTACCGGGCTCTCTGTGGCTTATGCGGGAAC	2526
242	eArgGlnAspProGlyAlaHisArgAsnLeuThrPnHisIleAlaArgLeu	259
2527	CAACCAAGACCCCGACAGAC.....CTGA	2552

```

259  rgvllleu.....SerProglyValThrPro..... 267
2553 AGCGCGTGGCGGGAGAGCCCGCGGAGTGGACAGTGGCGCGCCCGCCAGGC 2602
268 .....LeuAspAlaProCys.....CysLeuProglyLy 277
2603 TCGGGGGAATGTGTGTCACAGCATGCGCGCTGCGCCGTGCACCCAGAGCA 2652
277 sValThrLeuCysTrpGlnAlaProAspGln..... 287
2653 GCAAGAG...TCCTCGCGCGCGCCGCGAGCGCTCGGGCGTATCTCCGCA 2699
288 .....SerProCysGlnProLeuValProProAlaProGln 299
2700 CGACAGCGCCCGCTGGCGCCCTGCGACAGCGCTTGTGCGCGCCGCG..... 2743
300 LysAsnAlaThrValAsnGluProGlnAspPheGln..... 311
2744 .....CACTACTTCACAGGCGCTGTCGTGCA 2768
312 .....LeuValAlaGlyHisPro...AsnLeuCysValGlnValSerT 325
2770 CGCGTCGCAAGTTCAGAGCGCATCTCGAGGCGCTGTGCTGCTGACAGTGCCA 2819
325 hrTrpGluLysValGlnLeuGlnAlaCysSerTrpAlaAspSerLeuGly 341
2820 CCTAC.....GTGGACAGCTGTGACGCCGCTGGGGGC..... 2851
342 ProPheLysAspMetLeuLeuValGluMetLys...ThrGlyLeuAsnA 358
2852 .....CACTCGCGCAGTGTAGAGCGGCC 2874
358 snThrSerValCysAlaLeuGluProSerGlyCysThrPro..... 371
2875 GGACTTCTGTGC.....CCTTCCAGTGCCTGCGCCACAGCACT 2912
2913 ACGAGCTCGCGGTGACCTGTCGTGGAGCTCCGAGCGCTGTGCG 2960
seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:us-08-718-388-6
seq_documentation_block:
  Sequence 6, Application US/08718388
  Patent No. 6271362
  GENERAL INFORMATION:
    APPLICANT: MORIKAWA, MINORU
    APPLICANT: HARADA, NAOKI
    TITLE OF INVENTION: GENE ENCODING I9G FC REGION-BINDING
    TITLE OF INVENTION: PROTEIN
    NUMBER OF SEQUENCES: 29
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
      STREET: PO BOX 747
      CITY: FALLS CHURCH
      STATE: VA
      COUNTRY: USA
      ZIP: 22040-0747
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patent Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/718,388
      FILING DATE:
    CLASSIFICATION: 536
  ATTORNEY/AGENT INFORMATION:
    NAME: MURPHY JR, GERALD M
    REGISTRATION NUMBER: 28,977
    REFERENCE/DOCKET NUMBER: 0250-111
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (703) 205-8000

```


CITY: FALLS CHURCH
 STATE: VA
 COUNTRY: USA
 ZIP: 22040-0747
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0,
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/718,368
 FILING DATE:
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: MURPHY JR, GERALD M
 REGISTRATION NUMBER: 28,977
 REFERENCE/DOCKET NUMBER: 0230-111
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 205-8000
 TELEFAX: (703) 205-8050
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 16382 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 9..16223
 US-08-718-368-8

[illegible]

alignment_block:
US-09-899-471-2_COPY_21_440 x US-08-718-388-8

Align seg 1/1 to: US-08-718-388-8 from: 1 to: 16382

[illegible]

6281	...	AGGACACCTGGCGAGTACCTCTGTAGTGGACACCTGGACAGGACACCT	6327
131	alGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGluVal	147	
6328	TGGGGGCTGGAACACTTCACTGTACAGTGTAGCCAAATGACACCGGGGAGC	6377	
148	GlnIleTyrSerTyrThrLys.....ProArgTyrGlnLysG1	160	
6378	CAGGCTGTCACTACACACCGGACGTGTACACCTGGCAATCTACAAACACAG	6427	
160	uLeuAsnLeuThrGlnGlnLeuPro.....AspGlyAspAsnValL	174	
6428	CTTACACATAAGTGGCCCGCTGGCCCGGAAAGCTACAGTGTGGAGGGCTGT	6477	
174	eulLeuThrLeuAspValSerGlnGlnAspPheSerPheLeuTyr	190	
6478	TCGTCACTCTGT.....	6488	
191	LeuArgProValProAspAlaLeuLysSerLeuTyrTyrLysAsnLeuTh	207	
6489CCCTTCAGTGTAGACTGTCTCTGCACGACACACCTGAG	6528	
207	rgLProGlnAsnIleThrLeuAsnHisThrAspLeuValProCysLeuc	224	
6527	CGGCGCGGACGTGGGGTGTACCAACCTCAGGGCT.....	6563	
224	ysIleGlnValTyrSerLeuGlnProAspSerGlu.....	235	
6564TCGGTGGCTTTCGACGGGACACAGCTTCGTGGCTG	6599	
236	ArgVal.....GluPheCys.....	242	
6600	CGGGTGGCGGGCGGCTACCGGGCTCTGTGTGGCTTATCGGGAGACTA	6649	
242	eATsGlnAspProGlyAlaHisArgAsnLeuThrPheHisIleAlaArgLeuA	259	
6650	CAACCAAGACACCCGACAGCAGC.....CTGA	6675	
259	rgValLeu.....SerProGlyValTyrGln.....	267	
6676	AGGCGGTGGGGGGGAGAGCCCGCGGATGTGACGTGGCGCGGCCAGGGC	6725	
268LeuAspAlaProCys.....CysLeuProGlyLyu	277	
6726	TGGCGGGAATGTGTGTCCAAACCAATGCCCGTGGCCGTGACCCAGAGCA	6775	
277	sValThrLeuCysTyrPheAlaProAspGln.....	287	
6776	GCAAGAG.....TCCCTTGGCGGGCGGACGCTGGCGCGGTGATCTCCGCCA	6822	
288SerProCysGlnProLeuValProProValProGln	299	
6823	CGAGCGGCGCGCTGGCGGCTGTGCCAGCGGCTTGTGCGCGCCGCGC.....	6866	
300	LysAsnAlaThrValAsnGluProGlnAspPheGln.....	311	
6867CAGTACTTCCAGGGCTGCTTGTGTGA	6892	
312LeuValAlaGlnHisPro.....AsnLeuCysValGlnValSerT	325	
6893	CGCGTCCCAAGTTCAGGGCACTCTCGAAGGCGCTCTGTCTGTGCAAGTGCCA	6942	
325	hrTyrGlnLysValGlnLeuGlnAlaCysSerTyrAlaAspSerLeuGly	341	
6943	CCCTAC.....GTGGCAGCCTGTCAAGCCGCTGGGGCC.....	6974	
342	ProPheLysAspAspMetLeuLeuValGlnMetLys..ThrGlyLeuAsnA	358	
6975CAGCTCCGGAGATGGAGGGGGCC	6997	
358	snThrSerValCysAlaLeuGlnProSerGlyCysThrPro.....	371	
6998	GGACTTCTGTTC.....CCTTCCAGTGGCCCTGGCCACAGCAGCACT	7035	

21856 CACTGCCCG

alignment_block:

alignment_block:

Align seg 1/1 to: US-08-026-138E-7 from: 1 to: 3717

221	ProCysLeuCySLeIleGln.....	ValTrpSerLeuGlnProAs	233
	::::		
3062	CCCTGT...TGCACGCGACCTGCACCTACAGATCTCTCCCTCGAGACAG		3108
233	pSerGluArgValGluPheCysProPheArgGluAspProGly.....		247
	::::: ::: ::::	::::	
3109	AGGTACAGGGGCGCCATTCTCCCGCTATTCCCGAGACCCCGGAGCCGA		3158
248AlaHisArgAsnLeuTrpHisIleAlaArgLeuArg		259
	::::: ::: ::::	::::	
3159	CGACCTGCCGCTGCTCGGCGCGGACAGCTGCTCGCGGAGGCTGTGC		3208
260	ValLeuSerProGlyValTrpGlnLeu.....	AspAlaProCysCys	273
3209	TGCGCGCGGCTGGGCCAGGGCCCCGCCCTCGGACGCTTCCCTCC		3258
273	sLeuProGlyLysValThrLeuCySTrpGlnAlaProAspGlnSerPro.		289
	:::	:::	
3259	AGCTCCG.....	TGGCAGAAAGCTTCACTGATCCAA	3250
290CysGlnProLeuValProProValProGlnLysAsnAlaThrVal		304
	::::		
3291	CCCTGCTGCCGTCAGGTGTACCGGACAGCCT.....	GGCTGTGC	3331
305	AsnGluProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuCySva		321
	::::: :::		
3332	CATGTCCCAAA.....	GCGGCAATCT.	3355
321	IleValSerThrTrpGlnLysValGlnLeuAlaCys.....		334
	::::: ::::	:::::	
3356GCGGCACTGGGCTCAAAACAGAGTGTGGCGCTGCATCTACCG		3400
335SerTrpAlaAspSerLeuGly		341
	:::::	:::::	
3402	GGAGGCTGTGTGAGGCGCTGCCAGCAGGGGCTGGCGGCACCTGTGCAGC		3451
342	ProPheLysAspAspMetLeuLeuValGluMetLysThrGlyLeuAsnAs		358
3452	CCA.....GACACACTG.....		3463
358	nthrSerValCysAlaLeuGlnProSerGlyCysThrProLeuProSerM		375
	: :::	: :::	
3464TCTGCTCGCACACCATTACCCACT		3488
375	etAlaSerThrArgAlaAlaArgLeuGlyGluGlnLeuLeuGlnAsp.Ph		391
	:::	:	
3489	GCGGTTCY.....GCTGGGGGACTGTGCGCTCACCTCCACCCCTGT		3532
391	eArgSerHisGlnCysMetGlnLeuTrpAsn.....		401
	::: ::::		
3533	CCAGCA...CAGTCCCTGCTCATTTGAACTYGGAGCTCGAGCCTCATCAGAC		3579
401		401
3580	AGAGCGACGACCCCTGGGCTAGTACAGGCTACAGGACAGTGGGCTCT		3629
402AspAspAsnMetGlySerLeuTrpAlaCysProMetAspLysTrpI.		417
	:::::	:::::	
3630	AGAAGAGCTCAGCAGGGAAGCTTGTGGGACACAAAGGTTTCCAAAGTCT		3679
417	eHis 418		
3680	GCAC 3683		

seq_name: /cgn2_6/prodata/2/lna/6B_COMB.seq:US-09-041-886-24

Sequence 24, Application US/09041886
Patent No. 6235872
GENERAL INFORMATION:

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

```

```
alignment_scores:      Length: 464  
    Quality:          Gaps: 20  
    Ratio: 0.599  
Percent Similarity: 35.776 Percent Identity: 19.612
```

alignment_block:
US-09-899-471-2_COPY_21_440 x US-09-041-886-24 .

Align seg 1/1 to: US-09-041-886-24 from: 1 to: 4608

[illegible]

335 AG.....GCACCTTACCAATGT...GAGGCATCTTTAGGAGAT 369
146 GluValGlnIleThrPserTyrThr...LysProArgTyrGlnLysGluLe 161
:::||||| ||| ||| |||
370 TCGGCTCATATTATGTGACAGCAAAAGTTGCAGTAGACAGACCACT 419
161 uAsnLeuThrGlnGlnLeuProasp.....GlyAspAsnV 173
1:::||||| :::
420 GAGGTCCCTTTCACAGACAGAAATCTGTACAGCCCTTCATGGGAGACAG 469
173 allLeuThrLeuAspValSerGlnGlnIleAspPheSerPheLeuLeu 189
||||||| |||
470 TGTCTCAAGTGTGAAGTCATTGGGAG..... 498
190 TyrLeuArgProValProaspAlaLeuLysSerLeuThrPyrLysAsnLe 206
|||:::||||| ||| ||| ||| |||
499CCCATGCCAACAAATC.....CACTGGCAGAAAGAC.. 528
206 uThrGlyProGlnAsnIleThrLeuAsnHisThrAsp.....L 219
529CAACAAGACCTGACATCCCAATCCAGGTGACTCCCGAGTGGTG 571
219 euValProCysLeuCysIleGlnValThrSerLeuGluProaspSerGlu 235
:::||||| ::: ||| ||| :::
572 TCTTCCCTCTGAGAGCATTCAGATCAGCCGACTCCAAACGGGGGACATT 621
236 ArgValGluPheCysProPheArg..... 243
622 GGAATTACCGGATGCTCAGCTGAAATCCAGCCAGCTCAAGAACAGGAAA 671
244GluAspProGlyAlaHisArgAsnLeu 253
:::||||||| ||| ||| :::
672 TGAAGCAGAAATGACAAATTTTATCAGATCCAGACATGACATGACCTGT 721
253 rPhisIle..... 255
722 ATTTCTGCAAGACCAATCCAAATGTAGTACCATGAGGAAAAAGATGCT 771
255 255
772 GTCCGGAATGTGTGTTTCTGCTATCTCCACCAAGTTTACTGGTT 821
255 255
822 ACGAGGCGAGAAATCATCACTCAGTCTAAAAAGTATCTTTATTGG 871
255 255
872 GTGAGAGCACTGCTTATCTCCAAATGTGACAGATGATGACAGTGAATG 921
256AlaAr 257
922 TATACCTGTGTGTACATATATAAATGAAATATATAGTCCCTGTGAGA 971
257 gLeuArgValLeuSerProGlyValTyrGlnLeuAspAlaProCysCysL 274
:::||||| ||| ||| ||| |||
972 GCTCAGAGTCTTGGTTCG...CCATGGTTTAAATCATCTTCCACACC 1018
274 euProGlyLysValThrLeuCysTyrPglAlaProasp...GlnSerPro 289
|||:::||||| ||| :::
1019 TG.....TATGCTTATGAAAGCATGATGATTTGATTGAA 1053
290 CysGlnProLeuValProProValPro.....GlnLysAsnAl 302
||| ||| ||| ||| ||| |||
1054 TGTACAGTCTGTGAAAGCCTGTGCCCACTGGAATTTGATGAAAGATGG 1103
302 arhValAsnGluProGlnAspPheGlnLeuValAlaGlyHisProAsnL 319
:::||||| ::: ||| ||| ||| |||
1104 AGATGTGCTATCTTAGTATTTTACAGTAGTGGAGAGAGCAACT 1153
319 euCysValGlnValSerThrTrrPglLysValGlnLeuAlaCysSer 335
|||:::||||| ||| ||| ||| |||
1154 TAGGGATA..... 1161

336 TrpAlaAspSerLeuGlyProPheLysAspAsp.....MetLe 348
||||||| |||:::|||||
1162CTTGGGGTGTGAAAGTCAGATGAAGCTTTTATCAATG 1199
348 uLeuValGluMetLysThrGlyLeuAsnAsnThrSerValCysAlaLeuG 365
:::||||| ::: ||| ||| ||| |||
1200 TGTGGCTGAAATGAGCTGGAATGCCAGACAGTGCACAGCTCATTTG 1249
365 LndProSerGlyCysThrPro..... 371
|||:::||||| ||| ||| ||| |||
1250 TCCCTAAGCTGCAATCCCAAGTCCAGTGTCTCCCTTGGCTGCCAGA 1299
372LeuProSerMetAlaSerThrArgAlaAlaArgLeu 383
|||:::||||| ::: ||| ||| ||| |||
1300 GATGTGTCCCTGTGTGTTTCCAGCCGATTTGTCCCTCTC 1341

Sun Sep 29 09:32:47 2002

us-09-899-471-2_copy_21_440.rml

BASE COUNT ORIGIN	a	c	g	t	6 others
184	278	320	218		

Percent Similarity: 78.635 Percent Identity: 71.513

US-09-899-471-2_COPY_21_440 x BB609618

Align seg 1/1 to: BB609618 from: 1 to: 1006

563 TCCCAGGGTATGGCAGTAGATGCGCTTGCTGTCTGCCGGGCAAGT 612

959 TGC 961

seq_name: gb_est2:BG968034

```
seq_documentation_block:
```

DEFINITION	602832412F1 NCI_CGAP_CO24 Mus musculus cDNA clone IMAGE:4987035 5'
------------	--

ACCESSION	BG968034
VERSION	00000001

KEYWORDS	EST.	HOUSE	HOUSE
SOURCE			

ORGANISM

REFERENCE
1 (bases 1 to 784)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus

AUTHORS	TITLE
NIH-MGC http://mgc.ncl.nih.gov/ .	National Institutes of Health. Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Email: cyapus@email.mil.gov
 Tissue Procurement: Jeffrey E. Gre

CDNA Library Arrayed by: The T.M.A.C.E. Consortium (LNL)

Clone distribution: MGC clone distribution information can be obtained by: www.mgc.com.

<http://image.1ln1.gov>

High quality sequence stop: 646.

source	1. .784

```

/strain="FVB/N"
/lab="NIDDK-10000"

```

```
/clone="IMAGE:4987035"
/clone 14b="NCT CGAD C0341
```

```
/lab_host="DH10B (T1 phage-resistant)"
```


/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 173 a 231 c 229 g 151 t

alignment_scores:

Quality: 1128.00 Length: 208
Ratio: 5.502 Gaps: 1
Percent Similarity: 98.558 Percent Identity: 97.596

alignment_block:

US-09-899-471-2_COPY_21_440 x BG968034 ..

Align seg 1/1 to: BG968034 from: 1 to: 784

```

214 leuasnhrsthraspleuvalprocytleucysllleghvaltrpserle 230
|||||
3 ttaaacacacagacctgtgctccctgcactgcaggtgtgctgct 52
|||||
230 uglprowaspsergluarvalgluphbecsppropheargluasprow 247
|||||
53 agaccagacctctgagagggtcgaaattctgcccctccggagatcccg 102
|||||
247 lvalahlsatrgasnlleutrphtsllealargleuargvalleuserpro 263
|||||
103 gtgcacacagaaacctgtgcacatagccaggctggggctactgtcccca 152
|||||
264 glyalttrpglnleuaspalaprocytleuproglylvsvalthrlle 280
|||||
153 ggggtatggcagctagatgacgacctgtgctgtccggcgacagctaacact 202
|||||
280 ucysttrpglnalaproaspglinsertprocysglnproleuvalproproy 297
|||||
203 gtgctggcagcaccacacagacctgtgctgcagccacctgtggccacacag 252
|||||
297 alproglinsasnalathvalasngluuproglnasppheglnleuval 313
|||||
253 tgcgccagaagaacgccacctgtgatgagccacaaagatttccagttggctg 302
|||||
314 alaglylthsprowasnlleucysvalglnvalserthrtprglulysvalgl 330
|||||
303 gcagggccaccccaacctgtgtctcagctgagccacctgtggagaaaggttca 352
|||||
330 nleuglnalacysseertrpalaaspsertleuglyprophelysaspsam 347
|||||
353 gctgcagacctgtgtgctgactcctgtggggcccttcaagatgata 402
|||||
347 etleuvalglumetlsthrglyleuasnasnthrservalcysala 363
|||||
403 tgcgtgtagtagagatgaaacccggcctcaacacacatcactgtgtgcc 452
|||||
364 leuglnprosergllycsthproleuprosermetalaasertharagl 380
|||||
453 ttggaaacccagtgctgtacacacctgtccacagcatggactccacagagac 502
|||||
380 aalaargleuglygluileuvalleuaspheargserhsglncys 396
|||||
503 tgcctgcctggagagagagagttgctgcacgacttccgattccacacagtgta 552
|||||
397 metglnleutrpasnaspaspsammetllyserleutrpalaacysproke 413
|||||
553 atgcagctgtggaacgattgacacatgagatgcctatgatggcctgtcccat 602
|||||
413 taspbystrlllehlthararg 420
|||||
603 ggacaaagtacatccacagggcc 624
|||||

```

seq_name: gb_est2:BI332055

seq_documentation_block:

LOCUS BI332055 697 bp mRNA linear EST 30-JUL-2001
DEFINITION 602982878P1 NCI_CGAP_L19 Mus musculus CDNA clone IMAGE:5135755 5',
RNA sequence.
ACCESSION BI332055
VERSION BI332055.1 GI:15016712
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 697)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
<http://image.llnl.gov>
Plate: L1M11333 row: k column: 20
High quality sequence stop: 697.

FEATURES
SOURCE

1..697
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5135755"
/clone_id="NCI_CGAP_L19"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
141 a 212 c 205 g 139 t

alignment_scores:

Quality: 930.00 Length: 172
Ratio: 5.471 Gaps: 1
Percent Similarity: 98.837 Percent Identity: 98.256

alignment_block:

US-09-899-471-2_COPY_21_440 x BI332055 ..

Align seg 1/1 to: BI332055 from: 1 to: 697

```

250 argasnlleutrphtsllealargleuargvalleuserproglyvaltr 266
|||||
2 argasnlleutrphtsllealargleuargvalleuserproglyvaltr 266
|||||
266 pglinsuaspalaprocytleuproglylvsvalthrlleucysrrp 283
|||||
52 gcagctagatggcctgtgctgtgtgctggcgagcagctgacacagagatg 51
|||||
283 lnalaproaspglinsertprocysglnproleuvalproprovalprogl 299
|||||
102 aggcacacagacagagtgctgtgcagcactgtgtgacacagtgcccccag 151
|||||
300 lysasnalathvalasngluuproglnasppheglnleuvalaactg 316
|||||
152 aagaagccacactgtgataatgagccacaaagatttccatgtgtgtgagagcc 201
|||||
316 lsprowasnlleucysvalglnvalserthrtprglulysvalglnleugln 332
|||||
202 accccacctgtgtgtccaggtgagccacctgtggagaggttccactgcaaa 251
|||||
333 alacysseertrpalaaspsertleuglyprophelysaspsammetleule 349
|||||

```


252 GCGTCTGTGGGCTGACTCCTGGGGCCCTTCACAGATGATGCTGT 301
 349 VValGluMetLysThrGluLeuAsnAspThrSerValCysAlaLeuGluP 366
 302 AGTGGAGATGAACCGGCTCACACACATCAGCTGTGCTTGGAC 351
 366 roSerGlyCysThrProLeuProSerMetAlaSerThrArgAlaAlaArg 382
 352 CCACTGGCTGTACACACTGCCACCATGCGCTCCACGAGAGCTGCTGC 401
 383 LeuGlyGluGluLeuLeuGlnAspPheArgSerHisGlnCysMetGln 399
 402 CTGGAGAGGGGTTGCTCCAGACTTCGATCACAACACAGTATGACGCT 451
 399 UTTPAspAspAspAsnMetGlySerLeuTrpAlaCysPrometAlaPylr 416
 452 GTGAGACGATACACACATGGATGCTATGGGCTGGCCCATGGACAGT 501
 416 YrLLeuHisArgArg 420
 502 ACATCCACAGCGCG 515

seq_name: gb_est2:BI251117

seq_documentation_block:

LOCUS BI251117 839 bp mRNA linear EST 17-JUL-2001
 DEFINITION 602994315F1 NCL_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5150378 5',
 mRNA sequence.
 ACCESSION BI251117
 VERSION BI251117 GI:14800198
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 839)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.llnl.gov
 Plate: L1AM11371 row: m column: 03
 High quality sequence stop: 836.

FEATURES

SOURCE

Location/Qualifiers
 1..839
 /organism="Mus musculus"
 /strain="C57/B6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5150378"
 /clone_lib="NCL_CGAP_Mam5"
 /tissue_type="tumor, gross tissue"
 /dev_stage="7 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: Salt;
 Site:2: Not; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Robin Humphreys,
 NIH."

BASE COUNT 146 a 248 c 268 g 177 t
 ORIGIN

alignment_scores:
 Quality: 866.50 Length: 191
 Ratio: 5.009 Gaps: 2
 Percent Similarity: 90.576 Percent Identity: 90.576

alignment_block:

US-09-899-471-2_COPY_21_440 x BI251117

Align seg 1/1 to: BI251117 from: 1 to: 839

1 LeuGluArgLeuMetLupProGlnAspThrAlaArgCysSerLeuGlyLe 17
 259 CTGGAGAGACTATGAGCCTTCAGACACTGACAGCGCTGCTCTAGAGCCT 308
 17 uSerCysHisLeuTrpAspGlyAspValLeuGlySerLeuGly 34
 309 CTCCTGACCACTCTGGAGTGTACGTGCTGCTGCTGCTGGAAGCCTCC 358
 34 InSerAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeu 50
 359 AGTCGACCCAGGCGCCCTGTGCTAGTGGCTTACCCGCTGACAGAGCCTG 408
 51 ValLeuArgCysProGlnLysThrAspCysAlaLeuArgValVa 67
 409 GTGCTGAGGTGTCACAGAGACAGATTGGCCCTGTGCTGCTGCTGCT 458
 67 lValHisLeuAlaValHisGlyHisTrpAlaGluProGluGluAlaGlyL 84
 459 GGTCCACTGGCGGCTGATGGGCTGAGGAGAGAGCTGGAAGAGCTGGAA 508
 84 ySSerAspSerGluLeuGlnGluSerArgAsnAlaSerLeuGlnAlaGln 100
 509 AGTCGATTCAACATCCAGAGAGTCTAGAGACCCCTCTCCAGGCCAG 558
 101 ValValLeuSerPheGlnAlaTrpProIleAlaArgCysAlaLeuLeuG 117
 559 GTGCTGCTCTCTCTCCAGGCTTACCCATGCGCCGCTGCTGCTGGA 608
 117 VValGluValProAlaAspLeuValGlnProGlyGlnSerValGlySerA 134
 609 GGTCCAGGTGCGGCTGAGCTGAGTGGTGCAGCTGTGCTGCTGCTGCTG 658
 134 lValPheAspCysPheGlnAlaSerLeuGlyValAlaGlnIleTrp 150
 659 CGGTATTGACTGTTCCAGGCTAGCTTGGGGCTGAGTACAGTATGCTGG 708
 151 Ser.TyrThrLysProArgTyrGlnLysGluLeuAsnLeuThrGlnGln 167
 709 TCCCTAACACAGAGCCAGGTACAGAAAGAGCTCAACTCACACAGAGC 758
 167 eupProAsp..... 169
 759 TGCCGTGACTGCAGGCTCTGAAGTCCGGGACAGACATCCAGAGCTGTGG 808
 170 ...GlyAspAsnValLeuLeu 175
 809 AATGCTGACATGCTCTCTCTG 829

seq_name: gb_est2:BI765330

seq_documentation_block:

LOCUS BI765330 922 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603050641F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5150442 5',
 mRNA sequence.
 ACCESSION BI765330
 VERSION BI765330.1 GI:15756908
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 922)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M1476 row: b column: 11
 High quality sequence stop: 794.

FEATURES

source

Location/Qualifiers
 1. 922

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5190442"

/clone_id="NIH_MGC_116"

/lab_host="DH10B"

/note="Organ: pooled colon, kidney, stomach; Vector:
 pCMV-SPORT6; Site_1: NOTI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 26 yo male, 49 yo
 female, 71 yo male colon; 46 yo female kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb.
 Insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH_MGC library."

BASE COUNT 147 a 297 c 272 g 206 t
 ORIGIN

alignment_scores:
 Quality: 866.50 Length: 329
 Ratio: 3.610 Gaps: 8
 Percent Similarity: 72.948 Percent Identity: 57.751

alignment_block:

US-09-899-471-2_COPY_21_440 x B1765330

Align seg 1/1 to: B1765330 from: 1 to: 922

```

7 ProGlnAspThrAlaArgCysSerLeuGlyLeuSerCysHisLeuTrpAs 23
|||||.....|.....|.....|.....|.....|.....|.....|.....|
6 CCTAGGAGCGCTACCCACTGCTCTCGGCGCTCTCGGCGCTCTCGGCGG 55
23 pGlyAspValLeuCysLeuProGlySerLeuGlnSerAlaProGlyPro 40
|||||.....|.....|.....|.....|.....|.....|.....|.....|
56 CAGTACGACATACCTGCGCTGCGGCGGACATCGCTGCTCGGCGGCGG 105
40 AlLeuValProThrArgLeuGlnThrGluLeuValLeuArgCysProGln 56
|||||.....|.....|.....|.....|.....|.....|.....|.....|
106 TGCTGGCGCGCTACCGACCTCGACAGACAGAGCTGTGAGTGCCGAGAG 155
57 LysThrAspCysAlaLeuArgValArgValAlaHisLeuAlaValHis 73
|||||.....|.....|.....|.....|.....|.....|.....|.....|
156 GAGACCGACTGTGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCA 205
73 sGlyHisTrpAlaGluProGluGluAlaGlyLysSerAspSerGluLeu 90
|||||.....|.....|.....|.....|.....|.....|.....|.....|
206 T.....|.....|.....|.....|.....|.....|.....|.....|
90 InGluSerArgAsnAlaSerLeuGlnAlaGlnValValLeuSerPheGln 106
|||||.....|.....|.....|.....|.....|.....|.....|.....|
207 .....|.....|.....|.....|.....|.....|.....|.....|
107 AlaGlyProIleAlaArgCysAlaLeuLeuGluValGlnValProAlaAs 123
|||||.....|.....|.....|.....|.....|.....|.....|.....|
243 GCGTACCCCTACTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 292
123 pLeuValAlaProGlyLysValGlySerAlaValAlaPheAspCysPheG 140
|||||.....|.....|.....|.....|.....|.....|.....|.....|
293 CCGTTGCGACATTTGGTCAAGTGTGGGCTGTGTGTATATATGCTCTTC 342

```

```

140 LuAlaSerLeuGlyAlaGluValGlnIleTrpSerTyrThrLysProArg 156
|||||.....|.....|.....|.....|.....|.....|.....|.....|
343 AGCGTGGCGCTTAGGAGAGTACGAGTACGATCTGCTCTTACTACAGCCAG 392
157 TyrGlnLysGluLeuAsnLeuThrGlnGlnLeuPro.....|.....| 168
393 TACGAGAGAGAACTCAACACACACAGAGAGCTGCTGCGCTGCGCTGCTC 442
169 .....|.....|.....|.....|.....|.....|.....|.....|
443 AACGTGTACACAGATGCTGACACACGTCATCTGCTGCTGCTGCTGCTGCA 492
181 uGluGlnAlaPheSerPheLeuLeuTyrLeuArgProValProAlaAla 198
|||||.....|.....|.....|.....|.....|.....|.....|.....|
493 GAGGACGACCTTGGCGCTCTCCCTGTC.....|.....|.....|.....| 520
198 euLysSerLeuTrpTyrLysAsnLeuThrGlyProGlnAsnIleThrLeu 214
|||||.....|.....|.....|.....|.....|.....|.....|.....|
521 .....|.....|.....|.....|.....|.....|.....|.....|
215 AsnHisThrAspLeuValProCysLeuGlyIleGlnValTrpSerLeuG 231
|||||.....|.....|.....|.....|.....|.....|.....|.....|
554 AACGACACAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 603
231 LuProAspSerGluArgValGluPheCysProPheArgGluAspProGly 247
|||||.....|.....|.....|.....|.....|.....|.....|.....|
604 AACCTACTCCGTTAGGACACACATCTGCTGCTGCTGCTGCTGCTGCTG 653
248 AlaHisArgAsnLeuTrpHisIleAlaArgLeuArgValLeuSerProG 264
|||||.....|.....|.....|.....|.....|.....|.....|.....|
654 GCACACACAGAACTCTGCGCAAGCAGCGACCTGCGACCTGCGACCTGCA 702
264 yValTrpGlnLeuAspAlaProCysCysLeuProGlyLysValThrLeu 281
|||||.....|.....|.....|.....|.....|.....|.....|.....|
703 GAGGTGCTGCTGCGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 752
281 ySTrPAlaProAspGlnSerProCysGlnProLeuValProProVal 297
|||||.....|.....|.....|.....|.....|.....|.....|.....|
753 GCTGGCGGCTCGGCTGCGGAGACCTTCCGCGACCTG.....|.....| 796
298 ProGlnLysAsnAlaThrValAsnGluProGlnAsp.....PheGlnLeu 313
|||||.....|.....|.....|.....|.....|.....|.....|.....|
797 GTTACTCTGACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTGT 846
313 LuAlaGlyHisProAsnLeuCysValGlnValSer 324
|||||.....|.....|.....|.....|.....|.....|.....|.....|
847 TGAAGGACACCTAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 880

```

seq_name: gb_est1.A1463827

seq_documentation_block:

LOCUS A1463827 422 bp mRNA linear EST 09-MAR-1999
 DEFINITION V332a03.y1 Guaywoodford Beiler mouse kidney day 7 Mus musculus cDNA
 clone IMAGE:733036 5', mRNA sequence.

ACCESSION A1463827

VERSION A1463827.1 GI:4317857

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1999)
 Contact: Maria M/Washu-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800

[illegible]

/db_xref="taxon:10090"
 /clone="IMAGE:4946084"
 /clone_1lb="NCI_CAMP_Mam6"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: PCMV-SPORT6; Site.1: SalI;
 Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 library constructed by Life Technologies. Investigator
 providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 128 a 229 c 259 g 171 t
 ORIGIN

alignment_scores:
 Quality: 725.50 Length: 210
 Ratio: 4.218 Gaps: 9
 Percent Similarity: 81.905 Percent Identity: 78.095

alignment_block:
 US-09-899-471-2_COPY_21_440 x BG917539 ..

Align seg 1/1 to: BG917539 from: 1 to: 787

```

1 LeuGIuArgLeuMetGIuProGIuInASPThAlaArgCysSerLeuGlyLe 17
|||||
223 CTGGAGAGACTGATGGAGCCCTCAGAGACACTGCAGCGCTGCTCTTAGGCCCT 272
17 uSerCysHisLeuTPaspGlyAspValLeuCysLeuProGlySerLeug 34
|||||
273 CTGCTGCACACTCTGGGATGGTGACGTGCTCTGCTGCCTGAGAACCTCC 322
34 InSerAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeu 50
|||||
323 AGCTGTGCCCGAGGCCCTGTGCTAGTGCTTACCCTCCGTCACAGCGAGCTG 372
51 ValLeuArg_CysProGlnIuYThrAspCysAlaLeuArgValVal 67
|||||
373 GTGCTGAGGTGTGCCACAGAGACAGATTGGCCCTGTGTGCTGTGG 422
67 aLValHisLeuAlaValHisGlyHisTrpAlaGluProGluGluAla 82
|||||
423 TTGGTCCACTGTGGCCGTGCATGGGCGACTGGGCGAGACCTCGAAGAAGCT 472
83 GLYHisSerAspSerGluLeuGlnGluSerArgGlnAlaSerLeuGlnAl 99
|||||
473 GGAAGACTGTATTCAGACTCCAGACTCTAGGAAGCCCTCTCTCCAGGC 522
99 aGlnValValLeuSerPheGlnAlaTrpProIleAlaArgCysAlaLeu 116
|||||
523 CCAGGTGGTGCTCTCTCCAGGACATCCCAATGCCCGCGTGTCCCTGC 572
116 euGlnValGlnValProAlaAspLeuValGlnPheGlyGlnSerValG 132
|||||
573 TGGAGGTCCACAGGTGCCCGCTTGACCTGCTGACGCTGCTGCTGCGTGG 622
132 ySerAlaValPheAsp_CysPheGlnAlaSerPheGlyGlnGluValG 148
|||||
623 TTCTGCGGATTTGACTGTGTTGAGAGCTATGCTTGCGGCTGAGGGTAC 672
148 InLeuTrpSerTrpThrIuYsProArgTrpGlnGluLeuAsnLeuThr 164
|||||
673 AGATGTGGTCCCTACAGAGAGCCAGGCTACCAAGAAAGAGTCAACACTACA 722
165 GlnGlnLeuProAspGlyAspAsnValLeuLeuThrLeuAspValSerG 181
|||||
723 CACGACACTGGCTGACTGCAGGGGCTCT ..... 749
181 uGlnGlnAspPheSerPheLeuLeuTrpIleuArgProValProAspAla 198
|||||
750 ..... GAAGTCCGGGACACACA 765
  
```


198 eulysSerleutrP 202
 ::::|||||
 766 TCCAGAGCTGCTGG 779

seq_name: gb_est1:B1689749

seq_documentation_block:

LOCUS B1689749 829 bp mRNA linear EST 18-SEP-2001
 DEFINITION 603316255F1 NCL_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5356237 5',
 mRNA sequence.

ACCESSION B1689749
 VERSION B1689749.1 GI:15652378

KEYWORDS EST.
 SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 829)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.

CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: LLAM1905 row: n column: 14
 High quality sequence start: 4
 High quality sequence stop: 671.

Location/Qualifiers

FEATURES

source

1. 829

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone_image="5356237"

/clone_lib="NCL_CGAP_Mam6"

/sex="female, virgin"

/tissue_type="infiltrating ductal carcinoma"

/dev_stage="5 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI;
 Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 145 a 237 c 272 g 175 t

ORIGIN

alignment_scores:

Quality: 720.50 Length: 179

Ratio: 4.420 Gaps: 7

Percent Similarity: 91.061 Percent Identity: 87.709

alignment_block:

US-09-899-471-2_COPY_21_440 x B1689749

Align seg 1/1 to: B1689749 from: 1 to: 829

1 LeuGluArgLeuMetGluProGluAspThrAlaArgCysSerLeuGlyLe 17

246 CTGGAGAGACTGATGAGACCTCAGACACTGCACGCTCTCTAGGCTT 295

17 uSerCysHisLeuTrpAspGlyAspValLeuCysLeuProGlySerLeu 34

296 CTCTGACACCTCTGGGATGCTGAGCTGCTCTGCTGGAAGCTCC 345

34 InSerAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeu 50

346 AGCTTGCCCGGAGCCCTGTGTAGTCTACCCGCTGAGAGGAGCTG 395

51 ValLeuArgCysProGlnIysThrAspCysAlaLeuArgValArgVala 67

396 GTGCTGAGGTGTCCACAGAGACATGTCCTGTCTGCTGTGTGT 445

67 ValHisLeuAlaValHisGlyHisTrpAlaGluProGluGlnAlaGlyL 84

446 GTTCCACTTGCCGCTGCATGGCAGCTGGGAGAGCTGTAGAGAGCTGGA 495

84 ySerAspSerGluLeuGlnGlnSerArgAsnAlaSerLeuGlnAlaGln 100

496 AGTCTGATTCAGAACTCCAGAGCTTAGAGAACGCTCTCTCCAGGCCAG 545

101 ValValLeuSerPheGlnAlaTrpProIleAlaArgCysAlaLeuLeuG 117

546 GTGCTGCTCTCTCCAGGCTTACCCATCCGCGCTGTGTGCTGCTGG 595

117 ValValGlnValProAlaAspLeuValGlnProGlyGlnSerValGlySer 133

596 AGCTCAGAGTGGCCCGCTGACCTGTGACAGCTGTGCTGCTGCTGCT 645

134 AlaValPheAspCysPheGluAlaSerLeuGlyAlaGlnValGlnI 149

646 GCGGATTTGACTGTTTCAGAGCTTAGTCTTGCGGCTGAGAGGCTACAAA 695

149 LeuPheSerTyrThrIys.....ProArgTyrGlnIysGlnLeuAs 162

696 TCTGGTCTTACACAGAGGCCAGGGAACCAAGAAAGACTACAGCTCA 745

162 nLeuThrGlnGlnLeuProAspGly 170

746 AAACAGAGATGCTGACCTGAGAGG 770

seq_name: gb_est1:AA403935

seq_documentation_block:

LOCUS AA403935 397 bp mRNA linear EST 29-APR-1997
 DEFINITION Val2a03.r1 Guaywoodford Belier mouse kidney day 7 Mus musculus cDNA
 clone IMAGE:733036 5', mRNA sequence.

ACCESSION AA403935
 VERSION AA403935.1 GI:2057922

KEYWORDS EST.

ORGANISM house mouse.

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 397)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8581, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:450084

High quality sequence stop: 359.

Location/Qualifiers

1. 397

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone_image="733036"

/clone_lib="Guaywoodford Belier mouse kidney day 7"

/tissue_type="kidney"

/dev/strage="juvenile (7 days old)"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: Kidney; Vector: pInuScript SK-; Site_1:
 EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
 0190 dt. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
 -5' adaptor sequence: 5' GAATTCGACGAG 3' -3' adaptor
 sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3' Library provided
 by Lisa Guay-Woodford."

BASE COUNT 88 a 117 c 116 g 76 t
 ORIGIN

Alignment_scores:
 Quality: 706.00 Length: 132
 Ratio: 5.431 Gaps: 0
 Percent Similarity: 98.485 Percent Identity: 98.485

Alignment_block:
 US-09-899-471-2_COPY_21_440 x AA403935 ..

Align seg 1/1 to: AA403935 from: 1 to: 397

244 GLUAPPROGIALAHISARGASNEUTPHISILEALHARGLEUARGVA 260
 ||||||||||||||||||||||||||||||||||||||||||||||||
 3 GAAGATCCCGGTGCACACAGAACCTCTGGACATACCCAGCGCGGT 52
 ||||||||||||||||||||||||||||||||||||||||||||||||
 260 LLEUSERPROGLYVALTPGLNLEUASPALAPROCYSCYLEUPROGLYL 277
 ||||||||||||||||||||||||||||||||||||||||||||||||
 53 ACTGCCCCAGGGGTATGACAGTATGATGCGCTGCTGCTGCGGGGCA 102
 ||||||||||||||||||||||||||||||||||||||||||||||||
 277 YSVALTHLEUCYSTIRPGINALAPROASPINSERPROCYSGINPROLEU 293
 ||||||||||||||||||||||||||||||||||||||||||||||||
 103 AGGTACACTGTGTGGCGGACACGACGACGATCCCTCCAGCCACTT 152
 ||||||||||||||||||||||||||||||||||||||||||||||||
 294 VALPROVALPROGLNLYSASNALTHVALASGLUPROGLNAPPH 310
 ||||||||||||||||||||||||||||||||||||||||||||||||
 153 GTGCCACCAAGTCCCGCAGAGAACGCGACGTGAATGAGCCACAAGATT 202
 ||||||||||||||||||||||||||||||||||||||||||||||||
 310 EGLNLEUVALAAGLYHISPROASNLEUCYSVALGINVALSERTHTRPG 327
 ||||||||||||||||||||||||||||||||||||||||||||||||
 203 CCAGTTGGGCGAGGCGACCCCAACCTCTGTGTCCAGGTACGACCTGGG 252
 ||||||||||||||||||||||||||||||||||||||||||||||||
 327 LULYSVALGINLEUGINALACYSSETPALASPSEULEUGLYPROPH 343
 ||||||||||||||||||||||||||||||||||||||||||||||||
 253 AGAAGTTCAGCTGCACAGC TECTTGTGGGCTGACCTCTGGGCGCTTC 301
 ||||||||||||||||||||||||||||||||||||||||||||||||
 344 LYSASPAPMETLEULEUVALGILEUWETLYSTHGLYLEUASNAHTRSE 360
 ||||||||||||||||||||||||||||||||||||||||||||||||
 302 AAGGATGATATGCTGTACTGTAGTAGAAGAAACCGGCTCAACACATC 351
 ||||||||||||||||||||||||||||||||||||||||||||||||
 360 VALCYSLALEUGLUPROSERGLYCYSTHTRPROLEUPROSERMET 375
 ||||||||||||||||||||||||||||||||||||||||||||||||
 352 AGTGTGCTCTTGGAACCCAGTGGCTGTACACCACTGCCAGCATG 397
 ||||||||||||||||||||||||||||||||||||||||||||||||

Seq name: gb_esc2:BI770219

Seq_documentation_block:
 Locus BI770219 855 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603053267F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5203005 5',
 mRNA sequence.

Accession BI770219 GI:15761797

Keywords EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 855)
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM11508 row: m column: 22
 High quality sequence stop: 827.

FEATURES
 source Location/Qualifiers

1. 855
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5203005"
 /clone_lib="NIH_MGC_122"
 /lab_host="DH10B"
 /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
 Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
 anonymous pool of 24 week female lung, 16 week female
 spleen, and 20-22 week male spleens. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.4 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 026. Note:
 this is a NIH MGC Library."

BASE COUNT 147 a 268 c 265 g 175 t
 ORIGIN

alignment_scores:
 Quality: 697.00 Length: 225
 Ratio: 3.916 Gaps: 4
 Percent Similarity: 79.111 Percent Identity: 63.111

Alignment_block:
 US-09-899-471-2_COPY_21_440 x BI770219 ..

Align seg 1/1 to: BI770219 from: 1 to: 855

1 LEUGLARGLEUWETGLUPROGLNAPSPHRLAARGYCSERLEUGLYLE 17
 ||||||||||||||||||||||||||||||||||||||||||||||||
 214 CTGGAGAGGCTTGTGGGGCTCAGAGACGTACCCACTGCTCCGGGCT 263
 ||||||||||||||||||||||||||||||||||||||||||||||||
 17 USERCYHISLEUTRPAAPGLYASPVALLLEUCYSLEUPROGLYSER 34
 ||||||||||||||||||||||||||||||||||||||||||||||||
 264 CTCTGCGGCGCTTGGGACAGTACATCTGCTGCTGCGGGGACATG 313
 ||||||||||||||||||||||||||||||||||||||||||||||||
 34 INSERLAPROGLYPROVALLEUVALPROTHARGLEUGLTHRGLEU 50
 ||||||||||||||||||||||||||||||||||||||||||||||||
 314 TGCTGTGCTCGGGGCGGCTGCTGCGGCTTACGACCTCAGACAGATG 363
 ||||||||||||||||||||||||||||||||||||||||||||||||
 51 VALLEUARGYSPROGLNLYSTHASPYSALALEUARGVALARGVALA 67
 ||||||||||||||||||||||||||||||||||||||||||||||||
 364 GTGCTGAGGTGCGCAGAGAGACCGACTGTGACCTGTGCTGCGGTGC 413
 ||||||||||||||||||||||||||||||||||||||||||||||||
 67 IVALHISLEUVALAHISGLYHISTTPALAGLUPROGLYGLU..... 81
 ||||||||||||||||||||||||||||||||||||||||||||||||
 414 TGTCACCTTGCGCGCTGATGGGACGTGGAGAGAGCTCAAGCTGAGAA 463
 ||||||||||||||||||||||||||||||||||||||||||||||||
 82ALAGLYLSERASPSERGLUENGLUJLUSERATGASNAASER 96
 ||||||||||||||||||||||||||||||||||||||||||||||||
 464 AGTTTGGAGAGCAGCTGACTCAGGGGTGGAGAGACCTAGGATGCTCT 513
 ||||||||||||||||||||||||||||||||||||||||||||||||
 97 LEUGLALAGLINALVALLEUSERPHEGLNALATYRPROILEALARGCY 113
 ||||||||||||||||||||||||||||||||||||||||||||||||
 514 CTCAGAGCCCAAGTCGTCTCTTCCAGGCTACACCTACTGCGCGTG 563
 ||||||||||||||||||||||||||||||||||||||||||||||||
 113 SALALEULEUGLVALINALPROALASPLEUVALGINPROGLYGLNS 130
 ||||||||||||||||||||||||||||||||||||||||||||||||
 564 CGTCCTGCTGAGGTGCAGATGCTGCTGCTGTGATGATTTGGTCACT 613
 ||||||||||||||||||||||||||||||||||||||||||||||||
 130 ERTVALGISERLALVALPHASPCYPHEGLUALASERLEUGLYALAGLU 146
 ||||||||||||||||||||||||||||||||||||||||||||||||

10

COMMENT

Contact: Qinghua Zhang
 Shanghai Institute of Endocrinology, Rui-Jin Hospital
 197 Rui-Jin II Road, Shanghai 200025, P. R. China
 Tel: 86-21-64370045 (ex. 663332)
 Fax: 86-21-64743206
 Email: mbs@ems.stn.sh.cn
 This clone is available at Shanghai Hematology Institute in Shanghai.
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.
 Location/Qualifiers
 1. 843
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1lb="NPBHC06"
 /tissue_type="pituitary"
 /dev_stage="adult"
 /lab_host="SOLR"
 /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

FEATURES

source

ASE COUNT 150 a 243 c 268 g 181 t 1 others
 ORIGIN

Alignment scores:
 Quality: 693.50 Length: 298
 Ratio: 3.485 Gaps: 10
 Percent Similarity: 66.779 Percent Identity: 52.685

Alignment block:
 US-09-899-471-2_COPY_21_440 x AV752013 ..

Align seg 1/1 to: AV752013 from: 1 to: 843

```

1  LeuGluArgLeuMetGluProGlnAspTrpAlaArgCysSerLeuGlyLe 17
   |||||
16 CTGAGAGAGCTTGTGGGCTTCAGACGCTACCCACTGCTCCGGGCT 65
   |||||
17 userCysHisLeuTrpAspGlyAspValLeuGlyLeuProGlySerLeu 34
   |||||
66 CTCTGGCGGCTGTGGACAGTACATATCTGCTCCCTGGGGACATCG 115
   |||||
34 InsertAlaProGlyProValLeuValProThrArgLeuGlnThrGluLeu 50
   |||||
116 TGCGTCTCCGGGCCCGCTGCGGCGCTACGACCTGCAGACAGAGCTG 165
   |||||
51 ValLeuArgCysProGlnLysThrAspCysAlaLeuArgValArgVala 67
   |||||
166 GTGCTGAGGTCCAGAGAGACCGACTGTGACCTGTGCTGCGTGGC 215
   |||||
67 ValHisLeuAlaValHisGlyHisTrpAlaGluProGluGlu..... 81
   |||||
216 TGTCACACTGGCCGTCATGGCCACTGGGACAGCTGAAGAGAGAGAA 265
   |||||
82 .....AlaGlyLysSerAspSerGluLeuGlnGluSerArgAsnAlaSer 96
   |||||
266 AGTTGGAGGAGCAGCTGACTCAGGGGTGGAGAGCGCTAGGAATGCTCT 315
   |||||
97 LeuGlnAlaGlnValValLeuSerPheGlnAlaTrpProIleAlaArgCys 113
   |||||
316 CTCACAGGCCCAAGTCGTGCTCTCCCTTCAGGGCTACCCCTACTGCCGCTG 365
   |||||
113 ValLeuLeuGlnValValGlnValProAlaAspLeuValGlnProGlyGln 130
   |||||
366 CGTCTGCTGAGGTGCAAGTGCCTGCTGCTTGTGCGACTTTGGTCAGT 415
   |||||
130 erValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlnAlaGlu 146
   |||||
416 CTGGGGCTGTGTATATGACTCTTCGAGGCTGCGCTAGGAGTGAAG 465
   |||||
147 ValGlnIleTrpSerTrpThrLysProArgTrpGlnLysGluLeuAsnIle 163
   |||||

```

```

466 GTACGAACTGTCCTATTACTACAGCCAGGTACGAGAGAGAACATCAACCA 515
   |||||
163 uThrGlnGlnLeuPro.....AspGlyA 171
   |||||
516 CACACAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 565
   |||||
171 spAsnValLeuLeuThrLeuAspValSerGluGlnGlnAspPheSer.P 187
   |||||
566 ACAACCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 615
   |||||
187 heLeuLeuTrpLeuArg.....ProValProAspAlaLeuLysSer 200
   |||||
616 TCCTGATCTGATCAGTACGCTCCAGGGCCCAAAACCC..... 650
   |||||
201 LeuTrpTrpLysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisTrh 217
   |||||
651 .....GGGGGCCAA..... 659
   |||||
217 rAspLeuValProCysLeuGlyIleGlnValTrpSerLeuGluProAspS 234
   |||||
660 .....GAACTGATG 669
   |||||
234 erGluArgValGlu.....PheCysProPheArgGluAsp 245
   |||||
670 GACGGAATATTACTTGCACCAACAGCTGTTCTGGCCG..... 707
   |||||
246 ProGlyAlaHisArgAsnLeuTrpHisIleAlaArg..... 257
   |||||
708 .....GATCAGTGGGTGTGGACGATCCGTAGAGTCTTCCGGGAGC 748
   |||||
258 .....LeuArgValLeuSerProGlyValTrpGlnLeu 268
   |||||
749 GCGGCCCTTGAACGCGAGACGACTGACGTGTGGCAGCTG 788
   |||||

```

seq_name: gb_est2:BI820026

seq_documentation_block:

LOCUS BI820026 886 bp mRNA linear EST 04-OCT-2001

DEFINITION 603037207P1 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:5178409 5',

DEFINITION mRNA sequence.

ACCESSION BI820026

VERSION BI820026.1 GI:15931576

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 886)

AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

plate: L1444 row: m column: 02

High quality sequence stop: 884.

Location/Qualifiers

1. 886
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5178409"
 /clone_1lb="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Organ: pooled brain, lung, testis; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is

FEATURES

source

oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH-MGC Library.*

BASE COUNT 154 a 275 c 272 g 185 t

alignment_scores: Quality: 683.50 Length: 226
Ratio: 3.797 Gaps: 6
Percent Similarity: 79.646 Percent Identity: 63.274

alignment_block:

US-09-899-471-2_COPY_21_440 x B1820026 ..

Align seg 1/1 to: B1820026 from: 1 to: 886

```

1  LeuGluArgLeuMetGluProGlnAspThrAlaArgCysSerLeuGlyLe 17
247  CTGAGAGAGGCTTGCGGCTCAGGAGGCTACCCAGCTCTCCGGGCT 296
17  userCysHisLeuTrpAspGlyAspValLeuCysLeuProGlySerLeu 34
297  CTCTGCGCGCTCTGGAGAGTACATCTCTGCTGCTGGGAGACATCG 346
34  InsertAlaProGlyProValLeuValProThrArgLeuGlnThrGlu 50
347  TGCCCTGCTCCGGGCCCCGCGTGGGGCCCTACGACCTGCGACAGACTG 396
51  ValLeuArgCysProGlnLysThrAspCysAlaLeuArgVala 67
397  GTGCTGAGGTGCGACAGAGAGACCGACTGTGACTGTCTGCGTGGGC 446
67  1ValHisLeuAlaValHisGlyHisTrpAlaGluProGluGlu..... 81
447  TGTCACACTGGCGGTGACATGGGACACTGGAGAGCCCTGAAGATGAGGAA 496
82  ....AlaGlySerAspSerGluLeuGlnLeuSerAlaAsnAlaSer 96
497  AGTTTGGAGAGACACTGACTTAGGGGTGAGAGAGCCTAGCAATGCCCT 546
97  LeuGlnAlaGlnValValLeuSerPheGlnAlaTyrProIleAlaArg 113
547  CTCAGGCGCCCAAGTCTGCTCTTCCTCCAGGCTACCTACCTGCCCTG 596
113  sAlaLeuLeuGluValGlnValProAlaAspLeuValGlnProGlyGln 130
597  CGTCTGCTGGAGGTGAGAGTCTGCTGCTGCCCTTGTGCACTTGTGCTAG 646
130  erValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGlu 146
647  CTGAGGCTCTGTGTATATAGTCTGCTGAGGCTGCCCTCAGGAGAGAG 696
147  ValGlnIleTyrPseTyrThrLysProArgTyrGlnLysGluLeuAsn 163
697  GTAGAACTGTGCTCTATATCAAGCCAGGTACGAGAGAACTCAACCA 746
163  uThrGlnGln.....Leu.Pro.....AspGly 170
747  CACACAGCAGTGGCTTGGCTTCCCTGCTCAACGTGTTCAGCAGATGT 796
171  AspAsnValLeuLeuThrLeuAspValSerGluGlnAlaPheSerPhe 187
797  GACAACGTGACATCTGTCTGATATGCTCTGAGGAGACAGACATTCGG 846
187  eleu.LeuTyrLeuArgProValProAspAlaLeuLysSerLeuTyr 203
847  CTTCCTCTGTAC.....TGG 860
204  LysAsnLeuThrGlyProGlnAsn 211

```

861 AATGAGGTCCAGGCGCCCAAC 884

seq.name: gb_test2:BF791762

seq.documentation_block:

LOCUS BF791762

DEFINITION 602251953f1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4344301.5,

ACCESSION BF791762

VERSION BF791762.1 GI:12096816

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 946)

NIH-MGC http://imgc.ncl.nih.gov/.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

FEATURES

CDNA Library Preparation: Life Technologies, Inc.

source

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

Location/Qualifiers

1..946

Location/Qualifiers

High quality sequence stop: 680.

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

Location/Qualifiers

1..946

[illegible]

Align seg 1/1 to: AAS18132 from: 1 to: 2314

1 MetProValSerTrpPheLeuLeuSerLeuAlaLeuGlyValArgAsnProVal 17
199 ATCCCTGTCGTCCTGGTTCCTGCTGCTGCTGGACAGCTGGCGGAACCCGCT 248
17 ValAlaValSerLeuGluValArgLeuMetGluLeuProGlnAsnProThrAlaArgCys 34
249 GGGGTGTCCTCTGGAGAACATGATGAGGCTCAGAGACACAGCAAGCTGCT 298
34 exLeuGlyLeuSerCysHisLeuTrpAspGlyAspValLeuCysLeuPro 50
299 CTC TAGGCTCTCCCTGGCACACCTCTGGATGGTGAAGTGAAGTCTCTGCTGCT 348
51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuGly 67
349 GGAAGGCTCCAGTGTGCCCCAGGCGCTGTGCTAGTGGCTTACCCGCTGCA 398
67 nTrpGluLeuValLeuArgCysProGlnTrpThrAspCysAlaLeuArgVal 84
399 GACGGACCTGCTGCTGAGGTGGTCCACAGAGAACAGATGGCGCTTCGTG 448
84 aArgValValValHisLeuAlaValHisGlyHisTrpAlaGluProGlu 100
449 TCCGTGTGGTGGTCCACTGTGGCGCGGATGGGAGGAGGAGGAGGAGGCTGAA 498
101 GlnAlaGlySerSerAspSerGluLeuGlnGluSerArgAsnAlaSerLe 117
499 GAAGCTGGAAAGTCTGATTCAGAACMCTCCAGGAGCTTAGAGAACGCTCTCT 548
117 uGlnAlaGlnValValLeuSerPheGlnAlaValTrpValAlaArgCysAl 134
549 CCAGGCGCCAGGTGGTGGCTGCTCTCCAGGCGCTTACCCTCAGCCGCTGTG 598
134 ValLeuLeuGluValGlnValProAlaAspLeuValGlnProGlyGlnSer 150
599 CCTGTGTGGAGGTCCAGGTGCCGCTGTGACTGGTGCAGCGCTGTGCTCACTCC 648
151 ValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGluVal 167
649 GTGGGTCTCCGCGATTTGACTGCTTTCCAGGCTAGTCTTTGGGGCTGAGGT 698
167 GlnAlaTrpSerTrpThrTrpAspArgTrpGlnValGluLeuAsnLeuTr 184
699 ACAGACTGTGCTCTACACAGAACCCAGGTACACAGAAAGAGCTCAACTCA 748
184 hGlnGlnLeuProAspCysArgGlyLeuGlnValAlaArgAspSerTrpGln 200
749 CACAGCAGCTGCTGACTGACAGGGGCTTTGAAGTCCGGGACACACATCCAG 798
201 SerCysTrpAlaLeuProTrpLeuAsnValSerTrpAspGlyAspAsnVal 217
799 ACTGCTGTGGTCTCTGCGCTGCTCAATGTGTCTACAGATGGTGACATGT 848
217 IleuLeuThrLeuAspValSerGluGluGlnAspPheSerPheLeuLeuTr 234
849 CCTTCTGACACTGGATGTCTCTGAGGACACAGACCTTTACTCTTCACTGT 898
234 TrpLeuArgProValProAspAlaLeuLysSerLeuTrpTrpLysAsnLeu 250
899 ACCTGGGTCCAGTCCCGGATGCTCTCAATCTTGTGGTGAACAAACCTG 948
251 ThrGlyProGlnAsnIleThrLeuAsnHisThrAspLeuValProCysLe 267
949 ACTGGAGCTCAGAACATTTACTTTAAACACACAGACAGCTGTGCTCCGTCT 998
267 uCysIleGlnValAlaTrpSerLeuGluProAspSerGluArgValGlnPheC 284
999 CTGCATTCAGGTGTGGTGGCTAGAGCCAGACACTCTAGAGGGTGCATTTCT 1048
284 ySerProPheArgGluAspProGlyAlaHisArgAsnLeuTrpHisIleAla 300
1049 GCCCCTTCGGGGAAGATCCGGTCCACACAGGAACTCTCGGACACATAGCC 1098

301	ArgLeuArgValLeuSerProGlyValTrrpGlnLeuAspAlaProCysC	317
1099	AGGGTGGGGGTACTCTCCCGAGGGGTATGGACACTGATCGGCTTCTG	1148
317	SlrPProGlyLysValThrLeuCysTrpGlnAlaProAspGlnSerProC	334
1149	TCTGCGGGGCAGGTAACTACTGTGCTGGCAGGCACGACCAAGTCTCT	1198
334	ysgGlnProLeuValProProValProGlnLysAsnAlaThrValAsnGlu	350
1199	GCCAGGCCACTGTGTGCCACAGTGGCCCGACAGAAAGCCACTGTGAATAG	1248
351	ProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuCysValGlnVa	367
1249	CCAAAGATTTTCCACTGTGGTGGCAGGCCACCCCAACCTCTGTGTCCAGT	1298
367	lSerThrTrpGlyLysValGlnLeuGlnAlaCysSerTrpAlaAspSerL	384
1299	GAGCACTCGGGAGAAAGTTCACTGCTGCACCTGCACCTCTTGTGGGCTGACTTCT	1348
384	eugGlyProPheLysAspAspMetLeuLeuValGlnMetLysThrGlyLeu	400
1349	TGGGGCCCTTCAAGATATATAGCTGTATGTAGTGAATGAATAACGGGCTC	1398
401	AsnAsnThrSerValCysAlaAlaLeuGluProSerGlyCysThrProLeuP	417
1399	AACAAACATAGTCTGTGCTCTGGAAACCAAGGGGTGTACACCACTGGCC	1448
417	oSerMetAlaSerThrArgAlaAlaArgLeuGlyGlnGluLeuLeuGlnA	434
1449	CAGCATGGCGCTCCAGAGAGAGCTGTGCTTGGAGAGAGATTCTGTCAAG	1498
434	sPheAlaArgSerHisGlnCysMetGlnLeuTrpAsnAspAspAsnMetGly	450
1499	ACTTCCATCATCACCAAGTATGACGCTGTGGAAACATGCAACATGGGA	1548
451	SerLeuTrpAlaCysProMetAspLysTrpIleHisArgArgTrpValIle	467
1549	TTCGTATGGGCGCTGCCCATGATGACAACTACATCCACAGCGCTGGGTCTT	1598
467	vValTrpLeuAlaCysLeuLeuLeuAlaAlaAlaLeuPhePheLeuL	484
1599	AGATATGCTGGCTCTGCCTACTCTTGGCTGGCGGCTTCTTCTCTCTCC	1648
484	euleuLysLysAspArgArgLysAlaAlaArgLysSerArgTrpAlaLeu	500
1649	TTCTAAAAAGACCGCAGAAAGCGGCCCTGTGCTTCCGCAAGGGCTTGG	1698
501	LeuLeuHisSerAlaAspGlyAlaGlyTrpGluArgLeuValGlyAlaIle	517
1699	CTCTCTCACCTCCCGCAGCGAGCGGGCTACGAGCGCTGTGGAGACTT	1748
517	uAlaSerAlaLeuSerGlnMetProLeuArgValAlaValAspLeuTrpS	534
1749	GGGGTTCGGCGTTGAGCCAGATGCCACTGCGCGTGGCGGTGAGACTTGTGA	1798
534	erArgArgGlnLeuSerAlaHisGlyAlaLeuAlaSerPheHisHisGln	550
1799	GCGCGCGCGAGCTGAGCGCGCAGCGAGCGCTTACGCTGTTCCACACAG	1848
551	ArgArgArgIleLeuGlnGlnGlyLysValValIleLeuLeuPheSerPr	567
1849	CGAGCGCGGTATCTCGCAGAGAGGGTGGCTGTGAATCTTCTTCTTCGCC	1898
567	oAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnLeuGlnIleThrValGluP	584
1899	CGCGCGCGTGGCGAGTGTACGACGTGGCTGACGCTCCACAGACTGGAGC	1948
584	roGlyProHisAspAlaLeuAlaAlaTrpLeuSerCysValLeuProAsp	600
1949	CGGGCGCGCATGAGCCCTCCCGCGCTGGCTACGCTGCTGATACCCGAT	1998

601 PheLeuGlnGlyArgAlaThrGlyArgTyrValGlyValTyrPheAspG1 617
1999 TTCCTGCAAGGCGGCGGACCGGCTAGCTCGGGGTCTACTTGCAGCG 2048
617 YLeuLeuHisProAlaSerValProSerProPheArgValAlaProLeuP 634
2049 GGTGGTGCACCCAGACTGTGTGCTCCCGTCCCGTCCCGTCCCGTCT 2098
634 hSerLeuPProThrGlnLeuProAlaPheLeuAspAlaLeuGlnGlyG1 650
2099 TCTCCCTGCGCTCGACGCTCCGCTTCTGATGCATCTCAGGAGGCG 2148
651 CysSerThrSerAlaGlyArgProAlaAspArgValGlyArgValThrG1 667
2149 TCTCTCAGCTTCCGCGGCGGACCGGCGGCGGCGGCGGCGGCGGCGG 2198
667 nAlaLeuArgSerAlaLeuAspSerCysThrSerSerSerSerLualaProG 684
2199 GCGGCTGCGGCTCCGCGCTGACAGCTGTACTTCTAGCTCGGAGAGCCCG 2248
684 lYcysGcSgluGluTTPaspLeuGlyProCysThrThrLeuGlu 698
2249 GCTGCTGCAGAGAAATGGACCTGGACCTGCACTACACTAGAA 2292
seq.name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AA518133
seq_documentation_block:
ID AA518133 standard; cDNA; 2094 BP.
XX
XX AA518133;
XX
XX 26-MAR-2002 (first entry)
XX
XX Mouse DCRS7 reverse translation generic cDNA.
XX
XX
XX Mouse; DNA cytokine receptor subunit 7; DCRS7; phosphate labelling; ss;
XX gene therapy; protein therapy; immunological disorder.
XX
XX Mus musculus.
XX
XX M0200190358-A2.
XX
XX 29-NOV-2001.
XX
XX 23-MAY-2001; 2001WO-US16767.
XX
XX 24-MAY-2000; 2000US-206862P.
XX
XX (SCHE) SCHERING CORP.
XX
XX Gorman DM;
XX
XX WPI; 2002-106198/14.
XX
XX
XX Isolated antigenic human or mouse DNA receptor subunit-like
XX polypeptide useful for detecting antibodies generated in response to
XX presence of increased protein levels or immunological disorders -
XX
XX Disclosure; Page 20-21; 148pp; English.
XX
XX The invention relates to primate and rodent DNA cytokine receptor
XX subunit (DCRS) polypeptides and the polynucleotides encoding them. The
XX receptor, or their portions may be useful as phosphate labelling enzymes
XX to label general or specific substrates. The subunits may also be
XX functional immunogens to elicit recognising antibodies, or antigens
XX capable of binding antibodies. A combination, e.g., including a DCRS can
XX be used as an immunogen for the production of antisera or antibodies
XX capable of distinguishing between other cytokine receptor family members.
XX A purified DCRS can also be used as a reagent to detect antibodies
XX generated in response to the presence of elevated levels of expression,
XX or immunological disorders which lead to antibody production to the
XX translation generic cDNA.

XX
SQ Sequence 2094 BP; 247 A; 251 C; 393 G; 285 T; 918 other;
alignment_scores:
Quality: 2756.00 Length: 698
Ratio: 4.984 Gaps: 0
Percent Similarity: 79.226 Percent Identity: 71.490
alignment_block:
US-09-899-471-5 x AA518133
Align seq 1/1 to: AA518133 from: 1 to: 2094
1 MethProValSerThrPheLeuLeuSerLeuAlaLeuGlyArgAsnProVal 17
1 ATGCCNGTMSNTGGTYYTNTNTNMSNTYNGCNYTNGGNGMNAAYCCNGT 50
17 lValValSerLeuGlnArgLeuMetGluProGlnAspThrAlaArgCys 34
51 NGTNGTMSNTNGARBMGYTNAATGARGCCNARGAYACGCMNGTGTW 100
34 erLeuGlyLeuSerCysHisLeuTTPaspGlyAspValLeuCysLeuPro 50
101 SNYTNMGVYTWMSNTGYCAYTNTGGGAYGNGAGYCTNTNTGTYTNCN 150
51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuG1 67
151 GGNMSNTYNCARMSNGCNCNGCNGCNGTNTNGTCCNACNMGNTNTCA 200
67 nThrGluLeuValLeuArgCysProGlnLysThrAspCysAlaLeuArg 84
201 RACNGARTNGTNTNMGNTGTCCNARARACNGAYTGCCYNTGTG 250
84 alArgValValValHisLeuAlaValHisGlyHisThrAlaGluProGlu 100
251 TTMGNGTNGTNGTNCAYTNGCNGTNCAYGNGAYGCGCNGARGCNGAR 300
101 GluAlaGlyLysSerAspSerGluLeuGlnGluSerArgAsnAlaSerLe 117
301 GARGCNGNMAARMSNGAYTMSNGARYTNCARGARMSNGMNAAYCCNMSNT 350
117 uGlnAlaGlnValValLeuSerPheGlnAlaTyrProIleAlaArgCys 134
351 NCARGCNCARGTNGTNTNMSNTTTCARGCNTAYCCNATHGCMNGTGTG 400
134 lAlaLeuGluValGlnValProAlaAspLeuValGlnProGlyGlnSer 150
401 CNTNTNTGARGTNCARGTNCNGCNGAYTNTGNCARGCNGCNGARMSN 450
151 ValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGluVal 167
451 GTNGCNGMSNGCNGTNTTYGAYTGYTGARGCNGMSYNTNGCNGCARGT 500
167 lGlnIleThrSerTyrThrLysProArgTyrGlnLysGluLeuAsnLeu 184
501 NCARATHTGMSNTAYACNAAACNMGNTAYCARARGARYTNAAYTYNA 550
184 hGlnGlnLeuProAspCysArgGlyLeuGluValAlaArgSerIleGln 200
551 CNCARGARNTNCCNAYTCYMGNGGNTYTGAGTNGTNGMNGMNAATHCAR 600
201 SerCysTTPValLeuProTTPLeuAsnValSerThrAspGlyAspAsnVal 217
601 WSNRTGYGGTNTYTCNTGTGTYTNAAYTNWMSNACGAYGNGAYAYAG 650
217 lLeuLeuThrLeuAspValSerGlnGlnGlnAspPheSerPheLeuLeu 234
651 NYTNTNACNTNTNGAYTGNWMSNGARGARCARGAYTYTWSNTTYTNT 700
234 YrLeuAlaGProValProAspAlaLeuLysSerLeuTyrPyrLysAsnLeu 250
701 AYTYNMGCCGTCNCNCAATGCTTNAARMSNTNTGTATTAARAAYTNT 750

[illegible]

654 AGGAGAGAGGTAAGAACTGGTCTCTATTACTGACCCAGCCAGGTACGAGAAG 703

180 IuLeuAsnLeuThrGlnIleuProAspCysArgIleuGluValArg 196

704 AACCTAACCAACACACAGCAGCTGCTT 729

197 AspSerIleGlnSerCysTrpValLeuProTrpIleuAsnValSerThrAs 213

730 GCCCTGGCCCTGAGCTCAACGTGTACGAGA 758

213 pGjIAspAsnValLeuLeuThrIleuAspValSerGluGlnAspPhe 230

759 TGGTGACAACTGCATCTGCTTCTGAATCTCTCTGAGACAGCACTTCG 808

230 eRpHeLeuLeuTyrrLeuArgProValProAspAlaLeuLysSerLeuTrp 246

809 GCCTCTCCCTGTACTGCAATCAGTCCAGGGCCCCCAAAACCCGGTGG 858

247 TyrLysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAspLe 263

859 CACAAAAACCTGCATGCGACCGCAGATCATTAACCTTGAACACACAGACT 908

263 uValProCysLeuCysIleGlnValTrpSerLeuGluProAspSerGlu 280

909 GGATCCCTGCTCTGTATTCAGCTGTGGCTCGGCAACCTGACTCCGTTA 958

280 rGvalGluPheCysProPheArgGluAspProGlyIleHisArgAsnLeu 296

959 GGAACGACATCTGCCCTCTCAGGAGAGACCCCGCGCACACAGAACTTC 1008

297 TrpHisIleAlaArgLeuArgValLeuSerProGlyValTrpGlnLeuAs 313

1009 TGCCAACCCCGCCAGCTCGACTGTCGACCTGCAGACGCTGGCTGCTGA 1058

313 pAlaProCysCysLeuProGlyLysValThrLeuGlyTrpGlnAlaProA 330

1059 CGCACCGTGTCTGTCGCCGGAAGAAGCGGCACTGTCTGGCGGCTCCGG 1108

330 spInSerProCysGlnProLeuValProProValProGlnLysAsnAla 346

1109 GTGGGAGACCCCTCGCAGCAGCTAGTCCCAACCGCTTCTCTGGAGAACGTC 1158

347 ThrValAsnGluProGlnAspPheGlnLeuValAlaGlyHisProAsnLe 363

1159 ACGTGTGACAAAGTTCTCTGAGTTCACATGTGCAAGAGGCCACCTTAACCT 1208

363 uCysValGlnValSerThrTrpGluLysValGlnLeuGlnAlaCysSerT 380

1209 CTGTGTTCAGGTGAACACGCTGGGAAGAGTGCAGCTGCAGAGACTTGT 1258

380 rPAlaAspSerLeuGlyProPheLysAspAspMetLeuValGlnMet 396

1259 GGGCTGACTCTCTGGGGCTCTCAAAAGAGATGTGCTACTGTTGGACA 1308

397 LysThrGlyLeuAsnAsnThrSerValCysAlaIleGluProSerGlyC 413

1309 CGAGGGCCCCAGCAACACAGATCCCTGTGTGCTTGGAACCTCACTGGCTG 1358

413 sThrProLeuProSerMetAlaSerThrArgAlaIleArgLeuGlyGlu 430

1359 TACTTCTACTACCCGCAAAAGCTCCACGAGGGCAGCTGCTCTTGGAGAGT 1408

430 IuLeuLeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsnAsp 446

1409 ACTTACTACAAAGACTGTGAGTACGGCCAGTGTCTGACGACTATGG...GAC 1455

447 AspAsnMetCysSerLeuTrpAlaCysPrometAspLysTrpIleHisAr 463

1456 GATGACTTGGGAGCGCATATGGGCGTCCGCCAATGAGCAAAATACATCCACA 1505

463 gArGTrpValLeuValTrpLeuAlaCysLeuLeuLeuAlaIleAlaLeuP 480

1506 GGCCTGGGCCCCCTGCTGGCTGGGCTGCTCTACTCTTGTGGCGCTGGCTTT 1555


```

480 hepheheuleuleuLyAspArgArgIysAlaArgGlySer 496
      ::::::::::::::::::::::::::::
1556 CCCATCTCTCTTCACAAAAGATCACCGAAGCGCCGAGGCG 1605
      ::::::::::::::::::::::::::::
497 ArgThrAlaLeuLeuHisSerAlaAspGlyAlaGlyTrpGluArgLe 513
      ::::::::::::::::::::::::::::
1606 CGCGGGGGCTCTGCTCTACTACGCGAGTACTGGGTTTCAGAGCCT 1655
      ::::::::::::::::::::::::::::
513 uValGlyAlaLeuAlaSerAlaLeuSerGlnMetProLeuArgValAlav 530
      ::::::::::::::::::::::::::::
1656 GGTGGCGCCCTGCGCTGCGCTGCTGCGCTGCGCTGCGCTGCGCG 1705
      ::::::::::::::::::::::::::::
530 aAlaSerLeuPseArgArgGluLeuSerAlaHisGlyAlaLeuAlaTrp 546
      ::::::::::::::::::::::::::::
1706 TAGACCTGTGAAGCCGTGTAAGTACGACGCGGCGCTGCTGG 1755
      ::::::::::::::::::::::::::::
547 PheHisHisGlnArgArgArgIleLeuGlnGlyGlyValValIleLe 563
      ::::::::::::::::::::::::::::
1756 TTTCACGGGCGACGGGCGCACACCTGCGAGAGGCGCGCTGGTCTT 1805
      ::::::::::::::::::::::::::::
563 uLeuPseSerProAlaAlaValAlaGlnGlyGlnGlnTrpLeuGlnLeuG 580
      ::::::::::::::::::::::::::::
1806 GCTCTCTCTCCCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1855
      ::::::::::::::::::::::::::::
580 InThrValGluPro.....GlyProHisAspAlaLeuAlaAlaTrp 593
      ::::::::::::::::::::::::::::
1856 GGGTCTCCGGGCGCGGCGCGACGCGCGCGACGCGCTTCCGCGCTCG 1905
      ::::::::::::::::::::::::::::
594 LeuSerCysValLeuProAspPheLeuGlnGlyAlaGlnArgIleArgTy 610
      ::::::::::::::::::::::::::::
1906 CTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1955
      ::::::::::::::::::::::::::::
610 rValGlyValTyrrPheAspGlyLeuLeuHisProAspSerValProSerP 627
      ::::::::::::::::::::::::::::
1956 CGTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2005
      ::::::::::::::::::::::::::::
627 ropPheArgValAlaProLeuPheSerLeuProThrGlnLeuProAlaPhe 643
      ::::::::::::::::::::::::::::
2006 TTTTCCGACCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2055
      ::::::::::::::::::::::::::::
644 LeuAspAlaLeuGlnGlyGlyCysSerThrSerAlaGlyArgProAlaAs 660
      ::::::::::::::::::::::::::::
2056 CTGGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2105
      ::::::::::::::::::::::::::::
660 pArgValAlaGlyValThrGlnAlaLeuArgSerAlaLeuAspSerCys. 676
      ::::::::::::::::::::::::::::
2106 GAGAGCGGAGCAAGTGTGCTGCGGCGCTTCAAGCCAGCCGTGATAGTACT 2155
      ::::::::::::::::::::::::::::
677 .....ThrSerSerGluAlaProGly 684
      ::::::::::::::::::::::::::::
2156 TCCATCCCCCGGGGACCTCCCGCGCGCGGA 2184
      ::::::::::::::::::::::::::::
seq_name: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS46223
seq_documentation_block:
ID AAS46223 standard; cDNA: 2380 BP.
XX
XX AAS46223:
XX
XX
XX 18-DEC-2001 (first entry)
XX
XX Human DNA encoding PRO polypeptide sequence #299.
XX
XX PRO polypeptide; mammal; tumor; cancer; human; cattle; horse; sheep; ss;
XX dog; cat; pig; goat; rabbit; tumor necrosis factor alpha; TNF-alpha;
XX blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
XX adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
XX PCR primer.
XX Homo sapiens.
XX
XX MO200168848-A2.
XX

```

```

XX
PD 20-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06520.
XX
PR 01-MAR-2000; 2000WO-US05601.
XX
PR 02-MAR-2000; 2000WO-US05841.
XX
PR 03-MAR-2000; 2000US-187202P.
XX
PR 06-MAR-2000; 2000US-186968P.
XX
PR 14-MAR-2000; 2000US-189320P.
XX
PR 14-MAR-2000; 2000US-189328P.
XX
PR 15-MAR-2000; 2000WO-US06884.
XX
PR 21-MAR-2000; 2000US-190828P.
XX
PR 21-MAR-2000; 2000US-191007P.
XX
PR 21-MAR-2000; 2000US-191048P.
XX
PR 21-MAR-2000; 2000US-191314P.
XX
PR 28-MAR-2000; 2000US-192655P.
XX
PR 29-MAR-2000; 2000US-193032P.
XX
PR 29-MAR-2000; 2000US-193053P.
XX
PR 30-MAR-2000; 2000WO-US08439.
XX
PR 04-APR-2000; 2000US-194449P.
XX
PR 04-APR-2000; 2000US-194647P.
XX
PR 11-APR-2000; 2000US-195975P.
XX
PR 11-APR-2000; 2000US-196000P.
XX
PR 11-APR-2000; 2000US-196187P.
XX
PR 11-APR-2000; 2000US-196690P.
XX
PR 11-APR-2000; 2000US-196820P.
XX
PR 18-APR-2000; 2000US-198121P.
XX
PR 18-APR-2000; 2000US-198585P.
XX
PR 25-APR-2000; 2000US-199387P.
XX
PR 25-APR-2000; 2000US-199350P.
XX
PR 25-APR-2000; 2000US-199654P.
XX
PR 03-MAY-2000; 2000US-201516P.
XX
PR 17-MAY-2000; 2000WO-US13705.
XX
PR 22-MAY-2000; 2000WO-US14042.
XX
PR 30-MAY-2000; 2000WO-US14941.
XX
PR 02-JUN-2000; 2000WO-US15264.
XX
PR 05-JUN-2000; 2000US-209832P.
XX
PR 28-JUL-2000; 2000WO-US20710.
XX
PR 22-AUG-2000; 2000US-0644848.
XX
PR 24-AUG-2000; 2000WO-US23328.
XX
PR 08-NOV-2000; 2000WO-US30952.
XX
PR 01-DEC-2000; 2000WO-US32678.
XX
PR 20-DEC-2000; 2000WO-US34956.
XX
XX
XX (GENE ) GENENTECH INC.
XX
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
XX Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2001-602746/68.
XX
XX P-PSDB; AAU29322.
XX
XX
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
XX presence of tumours, such as prostate and breast tumours, in mammals and
XX to screen for modulators of the compounds
XX
XX Claim 2; Fig 597; 774pp; English.
XX
XX Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR
XX primers for PRO polypeptides of the invention. The sequences of the
XX invention can be used to detect the presence of a tumour in a mammal by
XX comparing the level of expression of a PRO polypeptide in a test sample
XX of cells from the animal and a control sample of normal cells, whereby a
XX higher level of expression in the test sample indicates the presence of a
XX tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,
XX pigs, goats and rabbits but are preferably human. The polypeptides can
XX be used to stimulate tumour necrosis factor (TNF) alpha release from human
XX blood, when contacted with it. A specific polypeptide can be used to
XX stimulate the proliferation or differentiation of chondrocyte cells. The
XX PRO proteins can be used to determine the presence of tumours and also
XX susceptibility to tumour development, particularly adrenal, lung, colon,
XX breast, prostate, rectal, cervical, or liver tumours, in mammalian
XX

```


34 erleuGlYleuSerCysHileutPaspGlyaspValleuCySleuPro 50
1188 GTGGGAGACCCTGCCAGCCACTGCTCCACCCGCTTCTCGGAGAACGTC 1237
333 CTCGGGCTCTCTCCGCTGAGACAGTACATCTGCTCCGCT 382
51 GlySerleuGlnSerAlaProGlyProValleuValProThrargleuG 67
383 GGGGACATCGTGCCTGCCCGGGCCGCTGCGGCGCTACGCACTGCA 432
67 nThrGluValleuValleuArgCysProGlnLysThrAspCysAlaLeuArgV 84
433 GACAGACTGCTGCTGAGTCCAGAGAGACACCACTGACCTCTGTC 482
84 alArgValValHisleuAlaValHisGlyHisThrAlaGluProGln 100
483 TGGCTGTGGCTGCTCCTGAGTGGCCATGGGCACTGGGAGAGCTTGAA 532
101 Glu.....AlaGlyLysSerAspSerGluLeuGlnGlnSer 113
533 GATCAGGAAAAATTGGAGAGACAGCTCAGCGAGGGGTGAGAGAGCTTGA 582
113 GAsnAlaSerleuGlnAlaGlnValValleuSerPheGlnAlaTyrProI 130
583 GAATGCTCTCTCCAGGCCCAAGTCGCTCTCTCCAGGCTTACCTTA 632
130 leAlaArgCysAlaLeuLeuGlnValGlnValProAlaAspLeuGln 146
633 CTCGCGGCTGCTGCTGCTGAGAGTCAAGTCCGCTGCTGCTGCTGCA 682
147 ProGlyGlnSerValGlySerAlaValPheAspCysPheGlnAlaSerle 163
683 TTGTGTGATCTGTGGGCTGTGTGTATATAGCTCTTGAGAGCTGCT 732
163 uGlyAlaGlnValGlnIleTyrPserTyrThrLysProArgTyrGlnLysG 180
733 AGGAGAGAGCTAGCAATCTGCTTATACCTACGACCCAGGTACGAGAA 782
180 luleuValleuThrGlnGlnLeuProAspCysArgGlyLeuGlnValArg 196
783 AACTCAACCAACACACAGAGCTGCT 808
197 AspSerIleGlnSerCysTyrPValleuProTyrPLeuAsnValSerThrAs 213
809GCCCTGGCTGGCTCAACGCTGCAGCA 837
213 pGlyAspAsnValleuLeuThrLeuAspValSerGlnGlnAspPheS 230
838 TGGTGAACAGCTGATCTGTGTGAATGCTCTGAGAGAGCACTTGC 887
230 erPheLeuLeuTyrLeuArgProValProAspAlaLeuLysSerLeuTyr 246
888 GCTCTTCCTGCTGATGAGATGAGTCCAGGGCCCCCAAAACCCGGGTG 937
247 TyrLysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAspLe 263
938 CACAAAACCTGACTGACGCGCAATCATTTACTTGAACACACAGACT 987
263 uValProCysLeuCysIleGlnValTyrPserLeuGlnProAspSerGlnA 280
988 GGTTCCTGCTGTATTCAGGTGAGTGGCTCTGGAACCTGACTCCGTTA 1037
280 rGlyAlaGluPheCysProPheArgGlnAspProGlyAlaHisArgAsnLeu 296
1038 GAGCAACATCTGCCCTTCAAGGAGAGCCCCCGCACACCAAGAACTTC 1087
297 TrpHisIleAlaArgLeuArgValLeuSerProGlyValTyrGlnLeuAs 313
1088 TGSCAAGCCGCGCGACTGCACTCTACCTGCAAGAGACTGCTCTGGA 1137
313 palArgCysCysLeuProGlyLysValThrLeuCysTyrPAlaProA 330
1138 CGACACGCTGCTGCTGCCGAGAAAGGAGACTGCTGCTGGGGGCTCGG 1187
330 spGlnSerProCysGlnProLeuValProProValProGlnLysAsnAla 346

1188 GTGGGAGACCCTGCCAGCCACTGCTCCACCCGCTTCTCGGAGAACGTC 1237
347 ThrValAsnGluProGlnAspPheGlnLeuValAlaGlyHisProAsnLe 363
1238 ACTGTGACAAAGCTTTCGAGTTCCTCATTCGTAAGAGCCACCTTAAC 1287
363 uCysValGlnValSerThrTyrPLeuLysValGlnLeuGlnAlaCysSer 380
1288 CTGTCTTCAGGTGAACAGCTCGGAGAAAGTGCAGCTGCAGAGAGCTTGT 1337
380 rPalAspSerleuGlyProPheLysAspAspMetLeuValGlnMet 396
1338 GGGCTGACTCCCTGGGCGCTTCMAAGACAGATGCTGACTGTGAGACA 1387
397 LysThrGlyLeuAsnHisThrSerValLysAlaLeuGlnProSerGlyC 413
1388 CGAGGCCCCAGACACACAGATCCCTGTGGCTTGGAAACCGAGTGGCTG 1437
413 sThrProLeuProSerMetAlaSerThrArgAlaAlaArgLeuGlyGln 430
1438 TACTTCACTACCCAGCAAAAGCTCCAGAGAGGAGCTGCGCTTGAGAGT 1487
430 luleuGlnAspPheArgSerHisGlnCysMetGlnLeuTyrPAsnAsp 446
1488 ACTTACTCAAAAGCTGACAGTCCAGGCTGCTGCTGAGCTATG...GAC 1534
447 AspAsnMetGlySerLeuThrPalAcysProMetAspLysTyrIleHisAr 463
1535 GATGACTTGGAGCGCTATGAGGCTCCCGCATGACAAATACATCCAA 1584
463 gArgTyrValleuValTyrPLeuAlaCysLeuLeuAlaAlaLeu 480
1585 GCGGTGGGCGCTCGTGTGCTGGCTGAGCTTCTTCCGCGGCTT 1634
480 hePhePheLeuLeuLysLysAspArgLys 491
1635 CCTCATCTCCCTTCMAAAAGATCAGCGCAAGGCTGAGAGCTC 1684
492AlaAlaArgGlySerArgThrAl 499
1685 TTGAACAGAGACTCCGCTCGGGGGCGGCGCCAGGGG...CGCGGGG 1731
499 aleuLeuLeuHisSerAlaAspValAlaGlyTyrGluArgLeuValGly 516
1732 TCTGCTCTCTACTCAGCGGATGAGCTGGGTTTCGAGCGCTGCTGGG 1781
516 lAleuAlaSerAlaLeuSerGlnMetProLeuArgValAlaValAspLeu 532
1782 CCTGCGCTGCGCTGTGCGACGCTCCGCTGCGCTGCGCTGAGACTG 1831
533 TrpSerArgArgGlnLeuSerAlaHisGlyAlaAlaAlaTyrPheHisH 549
1832 TGAAGCTCTGTAAGTGAAGCGCGAGCGGGCGCTGCTTGTTCACGC 1881
549 sGlnArgArgArgIleLeuGlnGlnGlyLysValValIleLeuLeuPheS 566
1882 GCAGGGGGCGGAGACCTCCAGAGAGCGGCGCTGCTGCTCTCTTCT 1931
566 erProAlaAlaValAlaGlnCysGlnIntyrPLeuGlnLeuGlnThrVal 582
1932 CTCCTGCGTGGCTGCGCTGCTGAGAGAGTGCATACAGATGGGCTGTC 1981
583 GluPro.....GlyProHisAspAlaLeuAlaAlaTyrPLeuSerC 596
1982 GGGGCGGGGGGCGACGCGCGACAGCGCTTCCGCGCTGCTGAGCTG 2031
596 sValLeuProAspPheLeuGlnGlyArgAlaThrGlyArgTyrValGly 613
2032 CGTGTGCTCCACTTCTTGCAGGGCGGGCGCGGCTGCTGCTGCTGCTG 2081
613 alTyrPheAspLysLeuLeuHisProAspSerValProSerProPheArg 629

230 erpHeLeuLeuTyrLeuArgProValPProAspAlaLeuLysSerLeuTrp 246
 888 GCGTCCTCCCTGACTGGAATCAGGTCCAGGGCCCCCAAAACCCCGGTGG 937
 247 TyrLysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAspLe 263
 938 CACAAAAACCTGACTGACCGCAGATCATTAATCTGAACACACAGACACT 987
 263 uValProCysLeuCysIleGlnValTrpSerLeuGluProAspSerLys 280
 988 GGTTCCTGCTGCTGATTCAGGTGTGGCTGTGAACCTGACTCCGCTTA 1037
 280 rGValGluPheCysProPheArgGluAspProGlyAlaHisArgAsnLeu 296
 1038 GGACACAAATGCCCCCTCAGGAGAGACCCCGCGCACACAGAACCTC 1087
 297 TrpHisIleAlaArgLeuArgValLeuSerProGlyValTrpGlnLeuAs 313
 1088 TGGCAGCGCCCGCGACTGCGACTGCTGACCTGCAGACTGGCTGCTGGA 1137
 313 PalaProCysCysLeuProGlyLysValThrLeuCysTrpGlnAlaPro 330
 1138 CGCACCGTGTGCTGCTGCGCGCAGAACGGGACACTGTGGCGGCTCCGG 1187
 330 spGlnSerProCysGlnProLeuValProProValProGlnLysAsnAla 346
 1188 GTGGGGACCCCTGCCAGCCACTGTGCCACCCGCTTCTTCTGGAGAACGTC 1237
 347 ThrValAsnGluProGlnAspPheGlnLeuValAlaGlyHisProAsnLe 363
 1238 ACTGGGACAAAGGTCTCTGAGTTCATTCATTCGAAAGCCACCTTAACCT 1287
 363 uCysValGlnValSerThrTrpGluLysValGlnLeuGlnAlaCysSer 380
 1288 CTGTGTAGGTGAGAACAGCTCGAGAACCTGCAGCTGCAGAGAGTCTTGT 1337
 380 rPalaSerSerLeuGlyProPheLysAspAspMetLeuLeuValGluMet 396
 1338 GGGCTGACTCCCTGGGGCTCTCAAGACGATGTGCTACTGTGGAACA 1387
 397 LysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluProSerGlyC 413
 1388 CGAGCCCCCAGACACAGATCCCTGTGCTGTGCTGTGACCCAGTGGCTG 1437
 413 sThrProLeuProSerMetAlaSerThrArgAlaAlaArgLeuGlyGlu 430
 1438 TACTTCATCTACCCAGCAAAAGCTCTCCAGAGGGCAGCTCGCTTGGAGAGT 1487
 430 LuLeuLeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsnAsp 446
 1488 ACTTACTACAAAGACTGCAGCTCAGGCCACTGTCTGCAGCATATG...GAC 1534
 447 AspAsnMetGlySerLeuTrpAlaCysPrometAspLysTrpHisArg 463
 1535 GATGACTGTGGGAGCTATGGGCTGCCCATGTGACAAATACATCCACAA 1584
 463 gArTrpValLeuValTrpLeuAlaCysLeuLeuAlaAlaLeuP 480
 1585 GCGCTGGGCGCTGTGTGGCTGCTGCTACTCTTGGCGGCTGCTTT 1634
 480 hepPheLeuLeuLeuLysLysAspArgArgLys... 491
 1635 CCCCATCTCTCTCTCAAAAAAGATCACCGCAAAAGCTGGCTGAGGCTC 1684
 492AlaAlaArgGlySerArgTrpAla 499
 1685 TTGAAMACGAGAGCTCCGCTCGGGGGCGCCGACGAGGCG...CGCGGCG 1731
 499 aleuLeuLeuHisSerAlaAspGlyAlaGlyTyrGluArgLeuValGly 516
 1732 TCTGTCTCTTACTACGACGAGACTCGGGTTTCGAGCGGCTGGTGGCG 1781
 516 lAleuAlaSerAlaLeuSerGlnMetProLeuArgValAlaValAspLeu 532

1782 CCGTGGCGTGGCGCTGTGCCAGCTGCGCTGCGGTGGCTGACACTG 1831
 533 TrpSerArgArgGluLeuSerAlaHisGlyAlaLeuAlaTrpPheHis 549
 1832 TGGAGCCGTGTGAATGAGCGCGAGGGGCCGTGCTGTGGTTTCACCC 1881
 549 sGlnArgArgArgIleLeuGlnGlyGlyValValIleLeuLeuPheS 566
 1882 GCAGCGCGCGCAGACCTGTCAGAGAGGGCGGCTGTGTGCTGCTTCT 1931
 566 erProAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnLeuGlnTrpVal 582
 1932 CTCGCGGTGGGTGGCTGTGCAGAGTGTGTACAGATGGGGTGTCC 1981
 583 GluPro.....GlyProHisAspAlaLeuAlaAlaTrpLeuSer 596
 1982 GGGCGCGGGCGCAGCGCCGACAGACGCTTCCGCGCTGCTGACAGTG 2031
 596 sValLeuProAspPheLeuGlnGlyArgAlaThrGlyArgValGly 613
 2032 CGTGTCTCCGACTTCTTGCAGGGCGGGCGCGGCGCAGCTACGTGGGG 2081
 613 aLtyrPheAspGlyLeuLeuHisProAspSerValProSerProPheArg 629
 2082 CTGTCTGTCAGAGCTGTCCACCGGAGCGCGTACCCGCTTCTTCCG 2131
 630 ValAlaProLeuPheSerLeuProThrGlnLeuProAlaPheLeuAsp 646
 2132 ACCGTGCGCTTCTTACACTGCTCCCACTGCGCAGACTTCTGGGGC 2181
 646 aleuGlnGlyLysSerThrSerAlaGlyArgProAlaAspArgValG 663
 2182 CTGTGACAGACTTGGCGCGCGCTTCCGGGGCTCCCAAGAGAGCGG 2231
 663 LuArgValTrpGlnAlaLeuArgSerAlaLeuAspSerCys.....Thr 677
 2232 AGCAAGTGTCCCGGCGCTTCCAGCAGCGCTGTGATAGCTACTTCATGCC 2281
 678 SerSerSerGluAlaProGly 684
 2282 CCGGGGACTCCCGCGCGGGA 2302
 seq_name: /sidsi/gcgdata/geneseq/geneseq-embL/NA2002.DAT.AAS18130
 seq_documentation_block:
 ID AAS18130 standard; cDNA; 2308 BP.
 XX
 AC AAS18130;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human DNAX cytokine receptor subunit 7 (DCRS7) cDNA.
 XX
 KW Human; DNAX cytokine receptor subunit 7; DCRS7; phosphate labelling; ss;
 KW gene therapy; protein therapy; immunological disorder.
 XX
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT CDS 181..2292
 FT /tag= a
 FT /product= "Human DCRS7"
 FT
 PN WO200190358-A2.
 XX
 PD 29-NOV-2001.
 XX
 XX 23-MAY-2001; 2001WO-US16767.
 XX
 XX 24-MAY-2000; 2000US-206862P.
 XX
 XX (SCHE) SCHERING CORP.


```

447 AspaSmMetGlySerLeuTriPalAcysPrometAspLysTyrIleHisAr 463
1477 GATGACTTGGAGCGCTATGGCGCTGCCCATGTGACAAATACATCCACA 1526
463 GArGTriValLeuValTriPLeuAlAcysLeuLeuAlaAlaLeuP 480
1527 GCGTGGGCGCCCTGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTT 1576
480 hepPheLeuLeuLeuLysLysAspArgLys..... 491
1577 CCCATCTCTCTCTCTCAAAAAGATCAGCGAAGGGTGGCTGAGGCTC 1626
492 .....AlaAlaArgLysArgTrrAl 499
1627 TTGAAGACGAGACGCTCGGCGGGGCGCGCGAGGGC...CGCGGGC 1673
499 aLeuLeuLeuHisSerAlaAspGlyAlaGlyTyrGluArgLeuValG 516
1674 TCTCTCTCTCTACTACGCGGATGACTCGGGTTTCAGACCGCTGCGG 1723
516 lalaAlaSerAlaSerGlnMetProLeuArgValAlaValAlaPLeu 532
1724 CCTGGGCTGCGCGCTGCTGCTGCGCTGCGCTGCGCTGCGCTGCTG 1773
533 TrpSerArgArgGluLeuSerAlaHisGlyAlaAlaAlaTrpPheHis 549
1774 TGGAGCGCTGCTGAGTACAGCCCGGCGGCGCGCTGCTGCTGCTG 1823
549 sGlaArgArgArgIleLeuGlnGlnGlyValValIleLeuLeuPhe 566
1824 GCAGCGGCGCGAGACCTGACAGAGGCGCGGCTGCTGCTGCTGCT 1873
566 erProAlaValAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 582
1874 CTCCCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1923
583 GluPro.....GlyProHisAspAlaLeuAlaAlaTrpLeuSer 596
1924 GGGCGCGGGGCGCGAGCGCGCGAGCAGCGCTTCCGCTGCTGCTG 1973
596 sValLeuProAspPheLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 613
1974 CGTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2023
613 aLTrpPheAspGlyLeuLeuHisProAspSerValProSerProPhe 629
2024 CTTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2073
630 ValAlaProLeuPheSerLeuProThrGlnLeuProAlaPheLeuAsp 646
2074 ACCGTGCGCGCTTACACTGCGCTGCGCTGCGCTGCGCTGCGCT 2123
646 aLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 663
2124 CTTGACAGCAGCTGCGCGCGCTGCTGCGCGCTGCTGCGCGCTG 2173
663 luaArgValThrlAlaAlaLeuArgSerAlaLeuAspSerCys.....Thr 677
2174 AGCAAGTGTCCCGGCGCTTACAGCAGCGCTGATGATGACTTCCAT 2223
678 SerSerSerGlnAlaProGly 684
2224 CCGGCGAGCAGCTGCGCGCGCTGCTGCGCGCTGCTGCGCGCTG 2244

```

seq_name: /STD1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF93875

seq_documentation_block:
 ID AAF93875 standard; cdna; 2499 BP.
 AC AAF93875;
 XX
 DT 23-MAY-2001 (first entry)

```

XX DE Human cDNA encoding a membrane or secretory protein clone PSEC0233.
XX KW Human; secretory protein; membrane protein; vaccine; gene therapy;
XX KW Rheumatoid arthritis; diabetes; ss.
XX OS Homo sapiens.
XX PN EP1067182-A2.
XX PD 10-JAN-2001.
XX PF 07-JUL-2000; 2000EP-0114090.
XX PR 08-JUL-1999; 99JP-0194179.
XX PR 11-JAN-2000; 2000JP-0118775.
XX PR 02-MAY-2000; 2000JP-0183766.
XX PA (HELI-) HELIX RES INST.
XX PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
XX DR WPI; 2001-093989/11.
XX DR P-PSDB; AAB88448.
XX PT Nucleic acids encoding secretory proteins/membrane proteins, useful in
XX gene therapy or as candidate target molecules in drug development -
XX PS Claim 1; SEQ ID 263; 609pp + CD ROM; English.
XX CC This invention relates to nucleic acid sequences AAF93744 - AAF93916
XX CC which encode human secretory or membrane proteins represented by
XX CC AAB88317 - AAB88419. Included in the invention are primers
XX CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
XX CC cDNA sequences of the invention. The invention also includes methods for
XX CC the production of antibodies directed against the proteins, and cDNA
XX CC sequences, which can be used in vaccines. The polynucleotide sequences
XX CC can be used in gene therapy. The polynucleotide sequences and the
XX CC proteins they encode may be used in the prevention, treatment and
XX CC diagnosis of diseases associated with inappropriate secretory
XX CC protein/membrane protein expression. The nucleic acids and complementary
XX CC sequences may also be used as DNA probes in diagnostic assays
XX CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
XX CC presence of similar nucleic acid sequences in samples. They may also be
XX CC used to study the expression and function of secretory proteins/membrane
XX CC polypeptides and their role in metabolism. The polypeptides may be used
XX CC as antigens in the production of antibodies against them and in assays to
XX CC identify modulators (agonists and antagonists) of expression and
XX CC activity. The antibodies and antagonists may also be used as therapeutic
XX CC agents to down regulate expression and activity. The antibodies may also
XX CC be used as diagnostic agents for detecting the presence of the
XX CC polypeptides in samples (e.g. by enzyme linked immunosorbent assay
XX CC (ELISA)). Examples of diseases which may be treated include rheumatoid
XX CC arthritis and diabetes.
XX SQ Sequence 2499 BP; 418 A; 821 C; 764 G; 456 T; 0 other;

```

alignment_scores:
 Quality: 2219.00 Length: 778
 Ratio: 3.977 Gaps: 10
 Percent Similarity: 71.722 Percent Identity: 57.841

alignment_block:
 US-09-899-471-5 x AAF93875 ..

Align seg 1/1 to: AAF93875 from: 1 to: 2499

```

1 MetProValSerTriPheLeuLeuSerLeuAlaLeuGlnGlnArgAspProVal 17
208 ATGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 257
17 lValValSerLeuGlnArgLeuMetGluProGlnAspThrAlaArgCys 34

```



```

1113 TCCCTCTTATTTGTCACAGACTGGACCGACGATCATTAACCTTGAC 1162
260 H1STHAspleuValProCysLeuCysIleGlnValTrpSerLeuGluPr 276
1163 CACACAGACTAGTCCCTGCTCTGTATTCAGGTGGCTGTGACCAACC 1212
276 OAspSerGluArgValGluPheCysProPheArgGluAspProGlyAlaH 293
1213 TGAAGCTTTAGGACGACATCTGCCCCCTTCAGGAGAGACCCCGCGCAC 1262
293 LAsrGAsnLeuTrpHisIleAlaArgLeuValValLeuSerProGlyVal 309
1263 ACCAGAACCTGTGGCAAGCCCGGAGTGGACATGCTGACCTTCAGAGC 1312
310 TrpGlnLeuAspAlaProCysCysLeuProGlyLysValTrpLeuCysTr 326
1313 TGGCTCTGAGACACACCTGCTGCTGCCCGCAGAACGGGACATGTGCTG 1362
326 PGIAlaProAspGlnSerProCysGlnProLeuValProProValProG 343
1363 GCGGCTCCGGGTGGGACCCCTGCCAGCCACGTGCTCCACCGCTTTCCT 1412
343 LAsrGAsnAlaTrpValAsnGluProGlnAspPheGlnLeuValAlaGly 359
1413 GGGAGAACCTGACTGTGGACAGAGTTCTCGAGTTCCATTGCTGAAGGC 1462
360 H1SPROAsnLeuCysValGlnValSerTrpTrpGluLysValGlnLeuG 376
1463 CACCTTAACCTGTGTGTTCAGGTGAACAGCTCGGAGAACTGACGTGCA 1512
376 nAlaCysSerTrpAlaAspSerLeuGlyProPheLysAspAspMetLeu 393
1513 GAGACTCTGTGGGTGACCTCCCTGGGCGCTTCMAAGACATGTGCTAC 1562
393 euValGluMetLysTrpGluLeuAsnAsnTrpSerValCysAlaLeuGlu 409
1563 TGTGGAGACAGACAGGCCCCAGACACAGATCCCTGTGCTTGACAA 1612
410 ProSerGlyCysTrpProLeuProSerMetAlaSerTrpArgAlaAlaAr 426
1613 CCAGTGGCTGTACTTACCTACCAACGACAAAGCCTCCAGC 1651
426 GLeuGlyGluGlnLeuGlnAspPheArgSerHisGlnCysMetGln 443
1652 .....C 1652
443 euTrpAsnAspAspAspMetGlySerLeuTrpAlaCysProMetAspLys 459
1653 TATGG...GACGATGACTTGGAGCGCTATGGGCTGCCATGACAA 1699
460 TyrlleHisArgAlaTrpValLeuValAlaTrpLeuAlaCysLeuLeuAl 476
1700 TGCATCCACAGAGGCTGGGCGCTGCTGTGCTGCGCTGCTACTCTTGGC 1749
476 aAlaAlaLeuPhePhePheLeuLeuLysLysAspArgArgLys.... 491
1750 CGCTGGCTTTCCTCATCTCTCTCTCAAAAAGATCAGGAAAGGAT 1799
492 .....AlaAlaArgGly 495
1800 GCGTAGGCTCTTGAACAAGAGCTCGCTGGGGGCGGCGCCAGGGGC 1849
496 SerArgTrpAlaLeuLeuLeuHisSerAlaAspGlyValaGlyTrpGluAr 512
1850 ...CGCGGCGCTGCTCTCTACTCATCGAGGATCGGGGTTTCGAGCG 1896
512 GLeuValGlyAlaLeuAlaSerAlaLeuSerGlnMetProLeuArgVala 529
1897 CCGTGGGGCGGCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1946
529 lAValAspLeuTrpSerArgArgGluLeuSerAlaHisGlyAlaLeuAla 545
1947 CCGTAGACTGTGAGACCGCTGCTGAACCTGAGCGCGGCGGCGCGCT 1996

```


[illegible]

PT New polypeptide encoding a human cytokine receptor Zcytoril4, for
treating inflammation e.g. rheumatoid arthritis -

PS XX Disclosure: Page 95-98; 112pp; English.

XX The invention provides a new human cytokine receptor designated Zcytoril4.
CC Zcytoril4 can be expressed by standard recombinant methodology. The
CC encoding nucleic acid is useful for detecting the expression of a
CC Zcytoril4 gene in a biological sample. Anti-Zcytoril4 antibodies can be
CC used to screen biological samples in vitro for the presence of Zcytoril4.
CC proteins, polypeptides and peptides having Zcytoril4 activity can be
CC administered to a subject who lacks an adequate amount of this
CC polypeptide, for treating inflammation and conditions such as Rheumatoid
CC arthritis. In contrast, Zcytoril4 antagonists (e.g. anti-Zcytoril4
CC antibodies) can be used to treat a subject who produces an excess of
CC Zcytoril4. Zcytoril4 nucleotide sequences can also be used to provide
CC Zcytoril4 to a subject. The present sequence represents a cDNA encoding
CC a variant of the human cytokine receptor Zcytoril4, designated Zcytoril4-1
CC The variant is a truncated form of the receptor polypeptide and lacks
CC amino acid residues 1-113 of Zcytoril4.
XX SQ

Sequence 1753 BP; 287 A; 584 C; 548 G; 334 T; 0 other;

alignment_scores:
Quality: 1888.50 Length: 594
Ratio: 4.053 Gaps: 7
Percent Similarity: 78.451 Percent Identity: 63.805

alignment_block:
US-09-899-471-5 x AAC85029 ..

Align seq 1/1 to: AAC85029 from: 1 to: 1753

110 GlngluserArgasnaAlaserleuglnalaglnvalValleuserPhegl 126
:::||||| ||||| ||||| :::::|||||
2 GAGGAGCCTAGGAATGCCTCTCCAGGCCCAAGTCGTCTCTCTCCGCA 51

126 nalatyrProIlealaIargcysAlaleuleuglnvalGlnvalProAlaA 143
||||| ||||| :::::|||||
52 GGCTTACCTACTGCGCGCTCGTGCTCTGGAGGTGCAAGTGGCCGTCG 101

143 spleuValGlnProglyInservAlgIservAlaValphaSpCySph 159
||||| ||||| :::::|||||
102 CCTTTGACATTGGTGTGCTGTGGGCTCTGTGTATATATACATCTCTTC 151

160 GluAlaserleuglnAlaglValglInllertPserTYrThrLysProAr 176
||||| :::::|||||
152 GAGGCTGCTTAGGAGAGTACAGTACAGATCTGTCTATATCTACAGCCAG 201

176 gTyGlnLysGluleuasnLeuthrGlnGlnLeuProaspCySarGlyYL 193
||||| :::::|||||
202 GTACGAGAAGAACTCACACACACAGAGCTGCCT 238

193 euGluValArgAspSerlleGlnserCysTrpValIleuProtrpleuan 209
:::||||| |||||
239 GCCCTGCCCTGGCTCAAC 256

210 ValSerThrAspGlyASPasnValleuLeuthrSAlasylSerGlnI 226
||||| :::::|||||
257 GTGTACGACAGATGTGACACAAGTGTGAGTCTGATATATCTCTGANGA 306

226 uGlnAspPheSerPheleuLeuTYrleuarProValProAspAlaleuL 243
||||| :::::|||||
307 GCAGCATCTGGCCCTCTCCCTACTAGTAATCAGATCCAGGGCCCCCAA 356

243 ysSerleuTrpTYrLysasnLeuthrGlyProGlnAsnIllethrLeuan 259
||||| :::::|||||
357 AACCCGGGTGGACAAAACCTGACTGAGCCACCCAGATCATATTACTTGNA 406

260 HisThrAspLeuValProCysLeucyslleGlnValThrpSerleuGuPr 276
||||| |||||
407 CACACAGACCTGGTTCCTGCTCTGTATATCAAGTGTGGCTCTGGAAC 456


```

562 eLeuLeuPheSerProAlaAlaValAlaGlnCysGlnIntrPLeuC
1300 CTTCCTCTCTCTCTCCGGGTGGGTGGCTGTGCAGCGAGTGGCTA
579 euGlnThrValGluPro.....GlyProHisAspAlaLeuAl
1350 ATGGGGTGTCCGGGGCCCGGGCGCAGCGCCCGCAGCGACTTCC
593 TrpLeuSerCysValLeuProAspPheLeuGlnGlyAlaGlnAlaThr
1400 TCGGTCAAGCTGCTGCTGCTCCGACTTCTTGCAGGGCGGGCGCCGAG
609 gTyTValGlyValTyTrPheAspGlyLeuLeuHisProAspSerValProS
1450 CTACGTGGGGGCGCTTGCAGAGGCTGTCCACCCGGAGCCGTAACCG
626 eTrpPheAlaGlyAlaProLeuPheSerLeuProThrGlnLeuProAla
1500 CCTTTTCCACCGCTCCGCTTTCACACTGACCTCCCACTGGCAAC
643 PheLeuAspAlaLeuGlnGlnGlyCysSerThrSerAlaGlyArgProAl
1550 TTCTCTGGGGCCCTCTGACAGCTCCGGCCCGCGCTCCGGGGCTGCCA
659 aAspArgValGluArgValThrGlnAlaLeuArgSerAlaLeuAspSerC
1600 AAGAGAACCGGACAGCAAGTGTCCCGGGCCCTTCAGCCAGCCCTGATGCT
676 ys.....ThrsSerSerGluAlaProGly 684
1650 ACTTCATCTCCCGGGGACTCCCGCCCGGGA 1681
seq_name: /SIBS1/gcdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ65269
seq_documentation_block:
ID AAZ65269 standard; DNA; 1752 BP.
XX
XX AAZ65269;
XX
XX 23-MAR-2000 (first entry)
XX
XX Human secreted protein gene 20.
XX
XX Human: secreted protein; cancer; tumour; developmental abnormality;
XX foetal deficiency; blood disorder; immune system disorder; inflammation;
XX autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
XX Schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
XX atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
XX digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
XX therapy; chromosome 3; ds.
XX
XX Homo sapiens.
XX
XX WO958660-A1.
XX
XX 18-NOV-1999.
XX
XX 06-MAY-1999; 99WO-0509847.
XX
XX
XX 12-MAY-1998; 98US-0085093.
XX 12-MAY-1998; 98US-0085094.
XX 12-MAY-1998; 98US-0085105.
XX 12-MAY-1998; 98US-0085180.
XX 18-MAY-1998; 98US-0085906.
XX 18-MAY-1998; 98US-0085920.
XX 18-MAY-1998; 98US-0085921.
XX 18-MAY-1998; 98US-0085922.
XX 18-MAY-1998; 98US-0085923.
XX 18-MAY-1998; 98US-0085924.
XX 18-MAY-1998; 98US-0085928.
XX 18-MAY-1998; 98US-0085928.
XX 18-MAY-1998; 98US-0085927.

```


PA (HOMA-) HUMAN GENOME SCI INC.

PI Ruben BM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
PI Olsen HS, Shi Y, Young PE, Wei F, Br wer LA, Soppet DR;
PI Lafleur DW, Endress GA, Ebner R;

DR WPI; 2000-062296/05.
DR P-PSDB; AAY76143.

PT New isolated human genes and the secreted polypeptides they encode
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders -
PT
PT

Claim 1; Page 308; 475pp; English.

AAZ65550 to AAZ65350 represent 97 isolated human secreted protein genes. This sequence was found to be present on human chromosome 3. AAY76124 to AAY76223 represent the secreted proteins encoded by the 97 human genes. The genes and their corresponding secreted polypeptides are useful for preventing, treating or ameliorating medical conditions, e.g. by protein or gene therapy. Also pathological conditions can be diagnosed by determining the amount of the new polypeptides in a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the 97 genes, based on which tissues they are most highly expressed in, and include developing products for the diagnosis or treatment of cancer, tumours, developmental abnormalities and foetal deficiencies, blood disorders, diseases of the immune system, autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney disorders, digestive/endocrine disorders, infections and AIDS. The polypeptides are also useful for identifying their binding partners. The sequences shown in AAY76224 to AAY76424 represent fragments of the secreted proteins.

Sequence 1752 BP; 324 A; 555 C; 522 G; 347 T; 4 other;

Alignment_scores:		
Quality:	1801.00	Length: 4966
Ratio:	4.248	Gaps: 4
Percent Similarity:	85.484	Percent Identity: 69.556

alignment_block:
US-09-899-471-5 x AAZ65269

Align seg 1/1 to: AA265269 from: 1 to: 1752

```

1 MetProValSerIrrPheLeuLeuSerLeuValLeuGluValArgAsnProVal 17
227 ATGCCCTGGCCCTGGATTCTTGCTCTCTGACAGTGGGGCCGAAGCCCAAT 276
17 lValValSerLeuGluValArgLeuMetGluProGlnAspThrAlaArgCys 34
277 GGTCCCTTCTCTGGAGAGGCTGTGTGGGGCCCTCAGAGACGCTAACCAATGCT 326
34 erLeuGluLeuSerCysHisLeuThrAspGluAspValLeuCysLeuPro 50
327 CTCGGGGCCCTCTCTGGCCGCGCTGTGGGACAGTACATATCTTGCGTGCCT 376
51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuGlu 67
377 GGGGAGCATCGTGCGCTCCGGGGCCCGTGTGGCGCCCTACGCACACTGCA 426
67 nThrGluLeuValLeuAlaArgCysProGlnIlyThrAspCysAlaLeuArg 84
427 GACACAGCTGGTCTCTGAGGTGCCAGAAAGGAACCGACATGTGACTCTGTTC 476
84 AlaArgValValHisLeuAlaValHisGlyHisIleThrAlaGluProGlu 100
477 TCGGTGTGGMTGTGCCACTTGGCGCTGCATGGGCACATGGGAAGACGCTGA 526
101 Glu.....AlaGlyLysSerAspSerGluLeuGlnGluSerAr 113

```

527 GATGAGGAAAAGTTTGAGAGGACCACTTACCTTAGGGGTGAGAGACCTTAG 576
113 gsnAlAserLeuGlnAlaGlnValValLeuSerPheGlnAlaIyrProI 130
577 GAATCCCTCTCTCCAGAGCCCAAGTGCCTCTCTCTCCAGAGCCCTACCTTA 626
130 leAlaIArgCysAlaLeuLeuGlnValGlnValProAlaSerPheValGln 146
627 CTGGCCGGCTGCTCTGCTGGAGGGTGAAGTCCCTGCTGCTCCCTTGTGAG 676
147 ProGlyGlnSerValGlySerAlaValPheAspCysPheGlnAlaSerLe 163
677 TTGGTGAAGTCGTGTGGGCTCTGTGGTATATAGACTCTCTTGAGCCCTCT 726
163 uGlyAlaGlnValGlnIleTyrSerTyrThrLysProArgTyrGlnLysG 180
727 AGGGAGTGAAGTACAAACTGGCTCTATACAGCCAGCCAGGTACGAGAGG 776
180 lndLeuAlnLeuThrGlnGlnIleuProAspCysArgGlyLeuGlnValArg 196
777 AATTCACACACACACAGAGAGCTGCTGACTGAGGGGCTGCACAACTGCG 826
197 AspSerIleGlnSerCysTyrPalaLeuProTyrPalaAsnValSerThrAs 213
827 AACACATCTCCAGATGCTGGGGCCCTGCCCTTGCTCAACGTTCCAGCAGA 876
213 pGlyAspAsnValLeuLeuThrLeuAspValSerGlnGlnAspPheS 230
877 TGGTGCAACAGCGATCGATCGTTCTGTAATGTCTCTGAGGAGCAGACTTCG 926
230 erPheLeuLeuTyrLeuArgProValProAspAlaLeuLysSerLeuTyr 246
927 GCCCTCTCCCTGACAGGAATGAGTCCAGGCCGCCCAAAACCCCGTGG 976
247 TyrLysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAspLe 263
977 CACAAAACCTACTGAGCCGAGATCATTACTTGAACACACAGACT 1022
263 uValProCysLeuCysIleGlnValTyrPserLeuGlnProAspSerGlnA 280
1027 GGTTCCTCCCTCTGTATATCAGGTGGGCTTCGTGAACTGACTCCGTTA 1076
280 rValAlaGlnPheCysProPheArgGlnAspProGlyAlaHisArgAsnLeu 296
1077 GACG.AMCATTCGCCCTTCAGAGMAGAACCCCGCCGACACCGAAGACCTC 1122
297 TrpHisIleAlaIArgLeuArgValLeuSerProGlyValTyrGlnLeuAs 313
1126 TGGCAAGCCGCCGAGCTGGACTGCTGACCTGCAGAGCTGGCTGCTGTA 1176
313 PalaProCysCysLeuProGlyLysValThrLeuYstrGlnAlaProA 330
1176 GCGACCTGCTCTGCTGCCCGCAGAAAGCGGCACTGTGTGGCGGCTCCGG 1222
330 spGlnSerProCysGlnProLeuValProProValProGlnLysAsnAla 346
1226 GTGGGAGACCCCTGCAGACCATGTGTCACCGCTGCTGCGGAGAAATGC 1276
347 ThrValAsnGlnProGlnAspPheGlnLeuValAlaGlyHisProAsnLe 363
1276 ACTGTGGACAAAGTTCTGTGAGATGCCATTTGCTGAAGAGCCACCTTAACCT 1322
363 uCysValGlnValSerThrProGlnLysValGlnLeuGlnAlaCysSerT 380
1326 CTGTCTTACAGTGAACAGCTCGGAAAGCTGAGCTGCAGAGAGTCTCTGT 1376
380 rPalaAspSerLeuAlaTyrProPheLysAspAspMetLeuLeuValGlnMet 396
1376 GGGCTGACTCCCTGGGGCTCTCCTCAAAGACAGATGTCTACTGTGGAGACA 1422
397 TyrThrGlyLeuAsnAsnThrSerValCysAlaLeuGlnProSerGlyC 413

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAC85028

AAC85028 standard; cDNA; 2076 BP

08-MAY-2001 (first entry)

Human cytokine receptor zcytor14 degenerate nucleotide sequence.

cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis; antiinflammatory; gene therapy; mice

[illegible]

Homo sapiens.

WO200104304-A1

18-JAN-2001.

30-JUN-2000; 2000WO-US18383

07-JUL-1999: 99TNS-03A885A

17VMO 1 2VWCEPUEPTEPOT TWO

100

XX

P-PSDB: AAB61880

Nov: 30
Dec: 31
Jan: 31
Feb: 28
Mar: 31
Apr: 30
May: 31
Jun: 30
Jul: 31
Aug: 31
Sep: 30
Oct: 31
Nov: 30
Dec: 31

treating inflammation e.g. rheumatoid arthritis -

Claim 8; Page 94-95; 112pp; English.

The invention provides a new human cytokine receptor designated zcytor14. zcytor14 can be expressed by standard recombinant methodology. The encoding nucleic acid is useful for detecting the expression of a zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be used to screen biological samples in vitro for the presence of zcytor14. Proteins, polypeptides and peptides having zcytor14 activity can be administered to a subject who lacks an adequate amount of this polypeptide, for treating inflammation and conditions such as rheumatoid arthritis. In contrast, zcytor14 antagonists (e.g., anti-zcytor14 antibodies) can be used to treat a subject who produces an excess of zcytor14. zcytor14 nucleotide sequences can also be used to provide zcytor14 to a subject. The present sequence represents the human cytokine receptor zcytor14 degenerate nucleotide sequence.

Sequence 2076 BP; 232 A; 273 C; 399 G; 265 T; 907 other;

alignment_block:
US-09-899-471-5 x AAC85028

Align seg 1/1 to: AAC85028 from: 1 to: 2076

1 MetPValSerTPhheuleuSerLeuAlalaGlyIarGlnProVa 17
1 ATGCCNGNCTCMTGGTTTNTNTNMSTNYTGCCYTTNGGNMGMSNCCNGT 50
17 ValValSerLeuGluArgLeuMetGluProGlnAspThrAlaArgCysS 34
51 NGTYNTMSTNYTGAMGVNVTGTCNGCNCNCARGAAGCCACNAOCATGYW 100
34 erLeuGlyLeuSerCysHisLeuTrpAspClyAspValLeuCysLeuPro 50
101 SNCCGNGNTNMSTNGCYMGTNTYTGGAAWMSNCAVATHYNTGYTMCCN 150
51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuGl 67
151 GGNNGAYATHGTCCNCCNCCNGCNCNGCNTYTTGCMCNACNCAATTNCA 200
67 nThrGluLeuValLeuArgCysSerProGlnLysThrAspCysAlaLeuArg 84
201 RACNCAARYTNCTNTNMGTGCARABARARACNGAATYGAGAYTTNTGY 250
84 alArgValValValHisLeuAlaValHisGlyHisTrpAlaGluProGlu 100
251 TTMGTGTCNGCTNCAATYTTNGCNGTNCAGCNCAATYGAGARGARCCGAR 300
101 Glu.....AlaGlySerAspSerGlnLeuGlnLeuSerAl 113
301 GAYGARGARAAATTTYGGCGCNCNCNGCAWMSNGSGTNCARARCCNMG 350
113 gAsnAlaSerLeuGlnAlaGlnValValLeuSerPheGlnAlaTyProI 130
351 NAAQCMNSNTNYTCARCGCACGACTGCTNTNMSTTYCARCGCTAACCA 400
130 LealArgCysAlaLeuLeuGlnGluValGluValProAlaAspLeuValGln 146
401 CNGCMGMGTGYCTNTNTNGARGTNCARGTCNCCNCCNGCNTNTNGTNCAR 450
147 ProGlnInsSerValGlySerAlaValPheAspCysPheGluAlaSerle 163
451 TTYYGNCARWMSGTNGCMSNCTGNTNTYGAATYGATYGARGCNCNTY 500
163 uGIValaGluValGlnIleTrpSerTyrThrIysProAlaGlyGlnLysG 180
501 NGCMNSNARCMGNALHTGTGMSATAACNARCCCMGTGATAGAARAAG 550
180 LuLeuAsnLeuThrGlnGlnLeuProAspCysArgGlyLeuGluValArg 196
551 ARYTNAAACAVACNARCARRYTNC.. 576
197 AspSerIleGlnSerCysTryValLeuProThrPheAsnValSerThrs 213
577GCNTYNTCCNTGGTYNAAAGTMMSNCCNGA 605
213 pGIYApsAsnValLeuLeuThrLeuAspValSerGluGlnGlnAspPheS 220
606 YGNGCATAYATGNCAYTNGTYNTMAATYTMNSNGARRCARCAATYTG 655
230 erPheLeuLeuTyrLeuArgProValProAspAlaLeuLysSerLeuTrp 246
656 GNYYTMSNYTATYAGGACARGTNCARCGNCCNCAARCCNMGTG 705
247 TyrIysAsnLeuThrGluProGlnAsnIleThrLeuAsnHISThrAspGle 263


```

706 CATAAARAYTTNACNGCNCARCTHATACNTYNAATCAATACNAAYTT 755
706 uvalProcysLeuGlyIleGlnValTrpSerLeuGluProAspSerGluA 280
756 NGNCCNTGTYTNTGATHCARGTGTGGCCTNTGARCNCNGAATWSGTNA 805
280 rylaIgluPheCysProPheArgGluAspProGlyAlaHisArgAsnLeu 296
806 GNAACNAAYATHGTGCCNTTYMGNGARGAYCCNMGCNCACARAAAYTN 855
297 TrpHisIleAlaArgLeuArgValLeuSerProGlyValTrpGlnLeuAs 313
856 TGGCARGCNCNCNTNTMGNTNTNTNACNTTNCARMSNTGTGTYTNTNA 905
313 palAProCysCysLeuProGlyLysValTrpLeuCysTrpGlnAlaPro 330
906 YGCNCNTGTYWSNTNCCNGCARGCNCNTNTGTYGCMGNCNCNG 955
330 spcInSerProCysGlnProLeuValProProValProGlnLysAsnAla 346
956 GNGCNGAYCCNTGTCARCNCNTNGCNCNCNTYTWMSNTGGARAAAYTN 1005
347 ThrValAsnGluProGlnAspPheGlnLeuValAlaGlyHisProAsnLe 363
1006 ACNGTNGATYARGTNTNGARTTTCCTNTNTNAARGNCATCCNAAYTT 1055
363 uCysValGlnValSerThrTrpGluLysValGlnLeuGlnAlaCysSer 380
1056 NTGTYGNCARGTNAAYWSNMSGARAAAYTNCARTNCARGARTGTGTYT 1105
380 rPalAspSerLeuGlyProPheLysAspAspMetLeuValGluMet 396
1106 GGGCNGAYWSNTNGCNCNTNTNARGAYGTYTNTYNTNTNGARACN 1155
397 LysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluProSerGlyCys 413
1156 MNGMGCCNCARGATYAAWMSNTYNTGCTNTNGARCCWMSGNGNTG 1205
413 sThrProLeuProSerMetAlaSerThrArgAlaAlaArgLeuGlyGlnG 430
1206 YACNMSGNTYTCMWSNARCCNMSGNCNCCNMGNNTNGNGART 1255
430 luleuLeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsnAsp 446
1256 AYTNTNCCARGAYTNCARMSNGCNCARTGTYTNCARTYNTGG...GAY 1302
447 AspAsnMetGlySerLeuTrpAlaCysProMetAspLysTrpHisArg 463
1303 GAYGAYTNTGNCNTNTGCGCTGTCNATGAYARATYATHCAYAA 1352
463 gArTrpValLeuValTrpLeuAlaCysLeuLeuAlaAlaLeuP 480
1353 RGNNTGGCCTNTGTYTNTGCTNTGTYTNTYTTGCGCNCNCNTYTW 1402
480 hepPheLeuLeuLeuLysLysAspArgLysAlaAlaArgLysSer 496
1403 SVTNTATHTYNTNTNAARAARGAYCAGCNAARGCNCNCNNGN 1452
497 ArgThrAlaLeuLeuHisSerAlaAspGlyAlaGlyTrpGluArgLe 513
1453 MCGCNCCTNTNTNTYTWMSGCNGAYWMSGNTTYGARMGNTY 1502
513 uValGlyAlaLeuAlaSerAlaLeuSerGlnMetProLeuArgValAla 530
1503 NCTNGCNCNTYTCMWSNCCNTNTGTCARTTNCCTNTNGNGTNGCNG 1552
530 alaAspLeuTrpSerArgGlyLeuLeuSerAlaHisGlyAlaLeuAlaTrp 546
1553 TNGAYTNTGWSNMGNCARGARTYTNMSGNCARGCNCCTGNCNTGG 1602
547 PheHisIleGlnArgArgGlyLeuGlnGlyGlyValValIleLe 563

```

```

1603 TTYCAYGNCNCCARMNMGNCARACNTNCCARGCNGCNGTNGTNTYT 1652
563 uLeuPheSerProAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnLeuG 580
1653 NTNTNTYWSNCCNGCNGCNCNTNTNTGTSNGARGTGTTCARGAYG 1702
580 lArThrValGluPro.....GlyProHisAspAlaLeuAlaAlaTrp 593
1703 GNGTWSNMGNCNCNGCNCARGCNCAYGNCAYGCTTNTMGNCNMSN 1752
594 LeuSerCysValLeuProAspPheLeuGlnGlyArgAlaThrGlyArgTy 610
1753 YTNWSNTGTYTNTNCCNGAYTNTYTNCCARGCNGCNCNGGWSMTA 1802
610 rValGlyValLTrpPheAspGlyLeuLeuHisProAspSerValProSerP 627
1803 YGTNGCNCNTGTYTGTGATYGNATNTNTNCAVCCNGAYGCTNGCNCNTY 1852
627 rOPheArgValAlaProLeuPheSerLeuProThrGlnLeuProAlaPhe 643
1853 TMTYTWGNACNCNTNGTNTTACNTNTNCCNMSGNCARTTNCARGATY 1902
644 LeuAspAlaLeuGlnGlyCysSerThrSerAlaGlyArgProAlaAs 660
1903 YTNMGNCNTYTNCCARCCNMGNCNCNMGNMSGNMGNTNTNCARGA 1952
660 pArGlyAlaArgValThrGlnAlaLeuArgSerAlaLeuAspSerCys. 676
1953 RMGNCNCARGARCTWSNMGNCNTNTNCARCCNCTNTNGAYWSNTAYT 2002
677 .....ThrSerSerGluAlaProGly 684
2003 TYCAYCCNCCNGNMACNCCNGCNCNCGN 2031

seq.name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: AAS18131
seq.documentatn_block:
ID AAS18131 standard; cDNA; 2109 BP.
XX
AC AAS18131:
XX
DT 26-MAR-2002 (first entry)
XX
DE Human DCRS7 reverse translation generic cDNA.
XX
KW Human; DNA cytokine receptor subunit 7; DCRS7; phosphate labelling; ss;
XX gene therapy; protein therapy; immunological disorder.
XX
OS Homo sapiens.
XX
PN M0200190358-A2.
XX
PD 29-NOV-2001.
XX
PF 23-MAY-2001; 2001WO-US16767.
XX
PR 24-MAY-2000; 2000US-206862P.
XX
PA (SCHE ) SCHERING CORP.
XX Gorman DM;
XX
DR WPI; 2002-106198/14.
XX
PT Isolated antigenic human or mouse DNA receptor subunit-like
PT polypeptide useful for detecting antibodies generated in response to
PT presence of increased protein levels or immunological disorders -
XX
PS Disclosure; Page 15-16; 148pp; English.
XX
CC The invention relates to primate and rodent DNA cytokine receptor
CC subunit (DCRS) polypeptides and the polynucleotides encoding them. The
CC receptors, or their portions may be useful as phosphate labelling enzymes

```


\$Q Sequence 2109 BP; 234 A; 267 C; 415 G; 269 T; 924 other;

```
Alignment_block:
US-09-899-471-5 x AAS18131 ..
```

Align seg 1/1 to: AAS18131 from: 1 to: 2109

```

601 AAYVSMATHCCMSMTGTGGCCNTTNCCTGTGGTTAAATGTMNSGCCGA 650
      ::::: ||| ::::::::::::::: ||||| ||||| ::::::: ||
213 pGLYAspaSnValLeuLeuThrLeuAspValSerGluGlnAspPheS 230
      ||| ||| ||| ||| ||| ::::: ||| ||| ||| ||| |||
651 YGGNGAYAAATGNCAYTGTGTYNTAAYGTNMSNGAGARCAATYTG 700
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
230 erPheLeuLeuTyrlLeuArgProValProAspAlaLeuLysSerLeuTrp 246
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
701 GNTYMSNTYNTAAYGAAYCARGTTCARGCCNCCNAACCNMGNTG 750
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 TyrlLysAsnLeuThrGlyProGlnAsnLleThrLeuAsnHsThrAspLe 263
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
751 CAYARAAYTNAAGGNCNCARATHATACAYTNAAYCAACGAYT 800
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
263 uValProCysLeuCysLleGlnValTrpSerLeuGluProAspSerGluA 280
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
801 NGTCCNCTGTYTNTGYAATHCARGTNTGGCCNTYNGARCCNAGVMSGTNM 850
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
280 rGValGluPheCysProPheArgGluAspProGlyAlaHsArgAsnLeu 296
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
851 GNCNAAATHTGYCCNTTYMGNGARGATCCNMGNCNCATCAARAAYTN 900
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
297 TrpHisLleAlaArgLeuArgValLeuSerProGlyValTrpGlnLeuAs 313
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
901 TGGCARGCNCNMGYNTMGNTYNTAATTCARMSNTGTYNTNGA 950
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
313 pAlaProCysCysLeuProGlyLysValThrLeuCysTrpGlnAlaProA 330
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
951 YGCNCCTGYMSNTTNCNCGCARGCCNCCNTYTGTCGMCNCNCG 1000
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
330 spGlnSerProCysGlnProLeuValProProValProGlnLysAsnAla 346
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1001 GNGNGNAYCCMTGYCARCCNTYNTGTCNCNCNTTMSNTGGARAAYGTN 1050
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
347 ThrValAsnGluProGlnAspPheGlnLeuValAlaGlyHisProAsnLe 363
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1051 ACNCTNGAY ..... 1055
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
363 uCysValGlnValSerThrTrpGluLysValGlnLeuGlnAlaCysSer 380
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1060 ..... GTNAAYMSNMSGARARARTTNCARYYTNCARGARTGYTNT 1095
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
380 rPalaAspSerLeuGlyProPheLysAspAspMetLeuLeuValGluMet 396
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1100 GGGCNGAYMSYNTTNGCCNCTNAARGAYGTYNTYNTYNTGARCN 1149
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
397 LysThrLysLeuAsnAsnThrSerValCysAlaLeuGluProSerGlyC 413
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1150 MGNGCNCNARGAYAAAMGMSNTTNTGYCXYTTCARGCCNMSGGWTG 1199
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
413 sThrProLeuProSerMetAlaSerThrArgAlaAlaArgLeuGlyGlu 430
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1200 YACNMSYTTNCNMSNARGCWMSNACNMGNCNMGTYTNGNART 1249
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
430 LuLeuLeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsnAsp 446
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1250 AAYTNTYNTCARGAYTTCARMSNCGNCARTGYTNTAAYTNTG...GAY 1296
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
447 AspAsnMetGlySerLeuTrpAlaCysPrometAspLysTrpLleHisAr 463
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1297 GATGAYTNGNCGNCTYNTGGCCNTGYCCNAAAGGAAARAAAYATHCCYAA 1346
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
463 gATGTrpValLeuValTrpLeuAlaCysLeuLeuLeuAlaAlaLeuP 480
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1347 RMGNTGGCCYNTNTNGTYNGCNGTGYTNTYNTTYTTCGNCGNCANTNM 1396
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
480 hePhePheLeuLeuLysLysAspArgGlyS ..... 491
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1397 SNTYNTAHTHYNTYNTAARAARGAYCAAGCNAARGNTGTYNTMGNTYNT 1446
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
492 ..... AlalaArgGlySerArghrAlaLe 500
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


1447 YTNARCARFARGAYGTNMGMSNGCNCNGCNCNGCNGMGNGMGNCAGNT 1496
500 uleuLeuHISserAlaaspGlyAlaGlyTYrGluArgLeuValGlyAlaL 517
1497 NYTNYTTAAVMSNGCNGAYGAYMSNGGNTTYGARMGAGNTNGNGCNGNY 1546
517 eulaSerAlaLeuSerGlnMetProLeuArgValAlaValaspLeuP 533
1547 TNGCNSMGNGCYTNTGYTCARTGCCNTNTMGNGTNGCNGTNAAYTNTGG 1596
534 SerArgArgGluLeuSerAlaHisGlyAlaLeuAlaTPrPheHisIsg1 550
1597 MSNMGNNGNGARNTYTNMSNGCNCARGNCNCNTNGCNTGTTCATGACNCA 1646
550 narGArgArgTLeuGlnGlnGlyValValIleLeuLeuPheSerP 567
1647 RMGNMGACARACNTTNCARGGNGGNGTNGTNTNTNTNTTYMSNC 1696
567 rolaAlaValAlaGlnCysGlnGlnTPrPleuGlnLeuGlnThrValGlu 583
1697 CNGNGCNGTNGCNYTNTGYMSNARGTGTYNCARGAYGNGTNGMSNGN 1746
584 Pro..... GlyProHisAspAlaLeuAlaIaTPrPleuSerCysVa 597
1747 CCGNGCNCNCAYGNCNCACATGACNTTYMGNGCNSMYTNMSNTGYGT 1796
597 lIeuproAspPheLeuGlnGlnIArgAlaThrGlyArgTYrValGlyValT 614
1797 NYTNCCNGAYTYYTNCARGNMGNCNCNGCNGMSNTAYTNGNGCNT 1846
614 yrPheAspGlyLeuLeuHisProAspSerValProSerProPheArgVal 630
1847 GYTTYGAYMGNTYTNCAAYCCNAYGNGTNCNGCNGCTNTTYMGNAON 1896
631 AlaProPheSerLeuProThrGlnLeuProAlaPheLeuAspAlaLe 647
1897 GTNCCNCTNTTYACNTYTNCCMSNCARTTCCNGATYYTYTNGMGCCNTT 1946
647 uGlnGlyGlyCysSerThrSerAlaGlyArgProAlaAspArgValGlu 664
1947 NCARCARCCNMGNGCNCNMGMMSGNGMNTTNCARGARMGNGCNGARC 1996
664 rGlyAlaThrGlnAlaLeuArgSerAlaLeuAspSerCys.....ThrSer 678
1997 ARGTNMSNMGNCYTNCCARCCNCTNTGAYMSNTAYTYCAYCCNCCN 2046
679 SerSerGluAlaProGly 684
2047 GGNACNMSNGCNCNGN 2064

eq. name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT :AAC85030

eq. documentation block:
D AAC85030 standard; cDNA; 1725 bp.
C AAC85030;
X
X 08-MAY-2001 (first entry)
X
E Human zcytor14 variant zcytor14-1 degenerate nucleotide sequence.
E
M Cytokine receptor: zcytor14; human; inflammation; rheumatoid arthritis;
M antiinflammatory; gene therapy; vaccine; variant; zcytor14-1; ss.
X
X Homo sapiens.
X
X WO200104304-A1.
X
X 18-JAN-2001.
D
D 30-JUN-2000; 2000WO-US18383.
F
F 07-JUL-1999; 99US-0348854.
R

PA (Zymo) ZYMOGENETICS INC.
XX Presnell SR, Burkhead SK, Powder SL;
PI
XX WPI: 2001-112618/12.
DR P-PSDB: AAB61881.
XX New polypeptide encoding a human cytokine receptor Zcytor14, for
PT treating inflammation e.g. Rheumatoid arthritis -
XX
PS Disclosure; Page 101, 112pp: English.

The invention provides a new human cytokine receptor designated Zcytor14. Zcytor14 can be expressed by standard recombinant methodology. The encoding nucleic acid is useful for detecting the expression of a Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be used to screen biological samples in vitro for the presence of Zcytor14 proteins, polypeptides and peptides having Zcytor14 activity can be administered to a subject who lacks an adequate amount of this polypeptide, for treating inflammation and conditions such as Rheumatoid arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14 antibodies) can be used to treat a subject who produces an excess of Zcytor14. Zcytor14 nucleotide sequences can also be used to provide Zcytor14 to a subject. The present sequence represents a degenerate nucleotide sequence of a variant of the human cytokine receptor Zcytor14 designated Zcytor14-1. The variant is a truncated form of the receptor polypeptide and lacks amino acid residues 1-113 of Zcytor14.

S0 Sequence 1725 BP; 192 A; 226 C; 332 G; 216 T; 759 other:

alignment_scores:
Quality: 1314.50 Length: 593
Ratio: 3.621 Gaps: 6
Percent Similarity: 61.214 Percent Identity: 45.025

alignment_block:
US-09-899-471-5 x AAC85030 ..

Align seg 1/1 to: AAC85030 from: 1 to: 1725

110 GlnGluSerArgAsnAlaSerLeuGlnAlaGlnValIleUserPheGl 126
:::||||| |||||::: |||||:|||||::: |||||
1 GARGARCCNNGNAAYGCNMSTNYTCARGCNCARGTNGTNMTWYSTVCA 50
126 nalaTyrrroIlealargcysAlaLeuleGluValGlnValProAlaa 143
||||| ||| |||::: |||||:|||||::: |||||
51 RGCTNTATYCACNCGNMGNTGYGTNNYTNGARGTCARCRTCCNGDNG 100
143 spleuValGlnProGlyGlnSerValGlySerAlaValPhaspcysPhe 159
||||| |||||::: |||||::: |||||::: |||||
101 CNTNGTNCARITYGNCARCMSNCTNGMNSNGTNGTATGATGYTTY 150
160 GluaIleSerLeuGlnAlaGlnValGlnIleTyrseryThrIlysProar 176
|||||::: |||||::: |||||::: |||||::: |||||
151 GARGCMCNCTNGCMWSNGARGTNGATHTFGMSNTFYAACARCCKNG 200
176 gTyrglInylsgluLeuasntleuthrGlnGlnLeubrOaspCysArngLYL 193
||||| ||||| ||| ||||| ||| |||
201 NTAYGARAAAGARTYNNAAYCACRCARCARYTMCCN..... 237
193 enGluValArgAspserIleGlnSerCysTrpValLeuProTyrlPeuasn 209
238 GCNTNCCNTGGTNNAY 255
210 ValSetThrAspGlyAspaSnValLeuLeuThrLeusPaValSergLgl 226
|||||::: |||||::: |||||::: |||||
256 GTMWSNCGNAGYGNGAAYGTNCAYTTNGTNTYNNAYGTNMSGARGA 305
226 uGlnAspPheSerPheLeuLeuTyrrLeuArgProValProAspalaleu 243
||||| |||||::: |||
306 RARCARVTYYTGNGTNNSTNTNTATYGAAGCAGTCAGTNCARGCNCNCA 355

530 alspLeuTPSerArgGluLeuSerLeuLanSgLyAlaLeuAlaTrp 546
1202 TNGAYTNTTGGMSMGMGNGARLTWNKSCNCARGGCNCNGTNGCTGG 1251
547 PheHisSgInrArgArgArgIleLeuGlnGlnGlyValValIleLeu 563
1252 TTYCAVGCNCARBMGNCNCARACQYTCARGAGGNGCNTNGTNGTNT 1301
563 uLeuPheSerProAlaIleValAlaGlnCysGlnGlnTrpLeuGlnLeuG 560
1302 MYTWTWTWMSNCNGCNGCNGTNGCNGTNTNTYWSMGARCTGGTNCARGAYG 1351
580 InrIleAlaGlnPro..... GlyProHisSgPalaeuAlaAlaTrp 593
1352 GNGTMSNGCNCNGCNGCNCATGNCNCAYAGCGCTTTCMGNCMSN 1401
594 LeuSerCysValLeuProAspPheLeuGlnGlyArgAlaThrGlyArgTy 610
1402 YTMSNTNGYGTNTTNCNGAYTYYTTCARGAGMGNGCNCNGCNGMSNTA 1451
610 rValGlyValTyTrpPheAspGlyLeuLeuHisProAspSerValProSerP 627
1452 YGTGNGCNGCTGTYTGAYMGNTTNTNCAYCCNGAYGCGTNGCNGCNY 1501
627 rPheArgValAlaProLeuPheSerLeuProThrGlnLeuProAlaPhe 643
1502 TMTTTCNMACNGTNCNGTNTTACATTCNCNCNCARFYTCNCNGAYTYY 1551
644 LeuAspAlaLeuGlnGlnGlyCysSerThrSerAlaGlyArgProAlaAs 660
1552 YTMGNCNCNTTCARCARCCNMGNCNCNMGMSNGMNGNTTNCARGA 1601
660 pArgValGluArgValThrGlnAlaLeuArgSerAlaLeuAspSerCys. 676
1602 RMGNGCNCARGTMSNMGNCNTTNCARCCNGCNGTNGAYTNTAYT 1651
677ThrSerSerGluAlaProGly 684
1652 TYCAYCCNCNGCNCNGCNGCNGNGN 1680
seq_name: /SIDS1/gcdata/geneseq/geneseqn_emb1/NA2001A.DAT:AAH99805
seq_documentation_block:
ID AAH99805 standard; cDNA; 1210 BP.
AC AAH99805;
XX
XX 16-OCT-2001 (first entry)
XX
XX
XX
DE Human protein encoding cDNA sequence SEQ ID NO:640.
XX
XX Human: cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antinflammatory; antithematic; antithrillitic; immunosuppressive;
KW antibacterial; endocrine; cardiac; central nervous system; virucide
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianemic; anem
KW antiagregant; haemostatic; vulnary; antitumor; osteopathic; eczem
KW dermatologic; antiallergic; antiasthmatic; antidiabetic; cytosatic
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infect
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflamat
KW antidiaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmun
KW genetic disease; haematopoietic disorder; platelet disorder; astma;
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder
KW neurological disorder; ss.
XX
XX Homo sapiens.
XX
XX MO200153455-A2.
XX
XX 26-JUL-2001.
XX

DM of: US-09-899-471-5 to: GenEmbl:* out_format : pfs

Date: Sep 29, 2002 1:05 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODEL=frim+g2n.model -DEV-xlh  
-O/cgn2_1/USPTO_spool/US09899471/runat.27092002.142439.17908/app_query.fasta.1.2519  
-DB=GenEmbl -OPT=fastap -SUPER=rge -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPT=0.000 -LOOPEXT=0.000 -GAPOP=4.500  
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELDP=6.000  
-DELXT=7.000 -START=1 -MATRIX=blissum62 -TRANS=human40.cdi  
-LIST=45 -DOCCALIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500  
-MINLEN=0 -MAXLEN=2000000000 -USER=US09899471.ecgn.1.11274  
-NCPU=6 -ICPU=3 -LONGLOG -DEV-TIMEOUT=120 -WARN-TIMEOUT=30  
-NO_XLPHY -WAIT -THREADS=1
```

Search information block:

```
Query: US-09-899-471-5  
Database: GenEmbl:*  
Database sequences: 1797656  
Database length: 187333701  
Search time (sec): 11857.370000
```

core list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
b.pat:AX350976	3723.00	4398.30	1.3e-236	2314	AX350976 Sequence 10 from Pat
b.pat:BC004759	3153.50	3725.17	3.9e-199	2015	BC004759 Mus musculus, clone M
b.pat:AX350978	2756.00	3254.14	6.8e-173	2094	AX350978 Sequence 12 from Pat
b.pat:AX073993	2368.50	2819.50	1.1e-148	2255	AX073993 Sequence 1 from Pat
b.pat:AX092430	2376.00	2803.04	9.1e-148	2380	AX092430 Sequence 161 from Pat
b.pat:AX180776	2376.00	2803.04	9.1e-148	2380	AX180776 Sequence 13 from Pat
b.pat:AX350973	2365.00	2790.28	4.7e-147	2308	AX350973 Sequence 7 from Pat
b.pat:BC006411	2233.00	2775.93	2.9e-145	2507	BC006411 Homo sapiens, clone M
b.pat:AX136341	2219.00	2616.69	2.2e-137	2499	AX136341 Sequence 263 from Pat
b.pat:AX073996	1888.50	2228.50	9.1e-116	1753	AX073996 Sequence 4 from Pat
b.pat:AX073995	1675.50	1974.77	1.2e-101	2076	AX073995 Sequence 3 from Pat
b.pat:AX350975	1645.00	1938.52	1.3e-99	2109	AX350975 Sequence 9 from Pat
b.pat:AX073998	1314.50	1548.95	6.5e-78	1725	AX073998 Sequence 6 from Pat
b.pat:AC024612	692.50	771.76	1.3e-34	166620	AC024612 Homo sapiens, clone M
b.pat:AC000783	692.50	771.66	1.3e-34	168665	AC000783 Homo sapiens, clone M
b.pat:AC022007	692.50	771.39	1.3e-34	173765	AC022007 Homo sapiens, clone M
b.pat:AC028196	692.50	770.51	1.5e-34	191834	AC028196 Homo sapiens, clone M
b.pat:AC018809	692.50	769.42	1.7e-34	216905	AC018809 Homo sapiens, clone M
b.pat:AC068099	583.00	642.59	2.0e-27	157773	AC068099 Homo sapiens, clone M
b.pat:HSN33055	581.00	688.13	5.8e-30	724	HSN33055 Homo sapiens, clone M
b.pat:AX136589	552.00	654.16	4.5e-28	694	AX136589 Sequence 511 from Pat
b.pat:AC026685	541.00	593.58	1.1e-24	145435	AC026685 Homo sapiens, clone M
b.pat:HSN33601	419.50	496.87	2.6e-19	726	HSN33601 Homo sapiens, clone M
b.pat:AC024612	362.50	381.00	7.4e-13	166620	AC024612 Homo sapiens, clone M
b.pat:HSN1106R	349.50	381.00	4.2e-15	319	HSN1106R Homo sapiens, clone M
b.pat:HSN335870	339.50	401.45	5.4e-14	734	HSN335870 Homo sapiens, clone M
b.pat:AX350982	337.50	378.86	9.7e-13	2012	AX350982 Sequence 16 from Pat
b.pat:AX456332	323.00	366.54	4.7e-12	4715	AX456332 Rattus norvegicus, clone M
b.pat:AX180778	317.00	365.88	5.1e-12	2138	AX180778 Sequence 15 from Pat
b.pat:AX136728	309.50	368.12	3.8e-12	613	AX136728 Sequence 650 from Pat
b.pat:AC068099	291.50	297.42	3.3e-08	157773	AC068099 Homo sapiens, clone M
b.pat:HSN335870	255.00	310.43	6.3e-09	284	HSN335870 Homo sapiens, clone M
b.pat:HSN335870	252.00	310.60	6.2e-09	187	HSN335870 Homo sapiens, clone M
b.pat:AC026196	208.00	196.80	0.0134	191834	AC026196 Homo sapiens, clone M
b.pat:AC096599	204.00	201.06	0.0078	69815	AC096599 Rattus norvegicus, clone M
b.pat:AC096599	203.00	199.88	0.0090	69815	AC096599 Rattus norvegicus, clone M
b.pat:HSN335870	189.00	234.43	0.0001	223	HSN335870 Homo sapiens, clone M
b.pat:AR034064	184.50	205.16	0.0046	3288	AR034064 Sequence 1 from Pat
b.pat:AR097654	184.50	205.16	0.0046	3288	AR097654 Sequence 1 from Pat
b.pat:AR097731	184.50	205.16	0.0046	3288	AR097731 Sequence 1 from Pat

gb.pat:AR102358	184.50	205.16	0.0046	3288	AR102358 Sequence 1 from P
gb.pat:AR104868	184.50	205.16	0.0046	3288	AR104868 Sequence 1 from P
gb.pat:AR131079	184.50	205.16	0.0046	3288	AR131079 Sequence 1 from P
gb.pat:AR137601	184.50	205.16	0.0046	3288	AR137601 Sequence 1 from P
gb.pat:MMU31993	184.50	205.16	0.0046	3288	MMU31993 Mus musculus, Interi

seq_name: gb.pat:AX350976

seq_documentation_block:

LOCUS	AX350976	2314 bp	DNA	linear	PAT 06-FEB-2002
DEFINITION	Sequence 10 from Patent W00190358.				
ACCESSION	AX350976				
VERSION	AX350976.1	GI:18616352			
KEYWORDS					
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (sites)				
AUTHORS	Gorman, D.M.				
TITLE	Mammalian receptor proteins: related reagents and methods				
JOURNAL	Patent: WO 0190358-A 10-29-NOV-2001;				
FEATURES	SCHERING CORPORATION (US)				
source	Location/Qualifiers				
	1..2314				
	/organism="unidentified"				
	/db_xref="taxon:32644"				
	/note="rodent; submitted Mus musculus"				
	199..2295				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD22836.1"				
	/db_xref="GI:18616353"				
	/translation="MRPVSFSLSLALGRPVVSLERLEPOTPARCSLGLSLMFG DYDLPGSLQSPGVLPTRLOTLELVLCRPORTDCLVRYVHLAVHGMPEEA GKSDSELDSESNASLQAOVLSFQAVPIARCALLEVQADLVQPGVSGAVDFCE ASLQAEVQIWSYTKRPEKELNLTQOLPCRGLEVDSDSCVLPVLANSTGDNVIL LTLVSEEDSFSLYLRPVPALSLMTKNTLNTHTDVLVCLCIOWSLP DSEVERECPREDPELHNLHRIARVSLPGVWOLADPCCLGKVTICWQAPDQSPC QPLVPVQKNAKVALREPDQFQVASHRNLCVYSTERYQOLACLNADSLGPFKDDML LVNKTGTLNNTSVCLBPSGCTPLPSMASTRARLEELLDERSHOCWOLMNDNG SLMACPMKRYHRRVNLVWACLALALFFLLKKDRKARGSTALLLSADA GYERLVAGLALSLQMPRLVAVDLMSRELSHGALAWHQRRIILOEGGVILLES PAVAACQOMLOLQVEPGRHDLALAMLSCLPDLQGRATGRVGYVFGGLADSV PSPPRVPLFSLPSQLPAPFALDAGGCSISAGRPADREVEVTAALSALDSISSRA PGCEEDNIGCTTLE"				
	259..2292				

CDS

mat_peptide	411 a	725 c	705 g	473 t
BASE COUNT	411 a	725 c	705 g	473 t
ORIGIN				

alignment_scores:

Quality: 3723.00	Length: 698
Ratio: 5.349	Gaps: 0
Percent Similarity: 99.713	Percent Identity: 99.570

alignment_block:

US-09-899-471-5 x AX350976

Align seg 1/1 to: AX350976 from: 1 to: 2314

1	MetProValSerTrpPheLeuSerLeuAlaLeuGlyAArgAsnProva	17
199	ATGCTGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	248
17	IValValSerLeuGluArgLeuMetGluProGlnAspThrAlaArgCys	34
249	GATGCTCTCTCTGAGAGACTGATGAGCTGAGAGACTGAGCTGCT	298
34	erLeuGlyLeuSerCysHisLeuTrpAspGlyAspValLeuCysLeuPro	50
299	CTCTAGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	348

51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuG 67
 349 GGAAGCCTCCAGTCTGCCCCAGGCCCTGTGCTAGTCCCTACCCGCTGCA 398
 67 nThrGluLeuValLeuArgCysProGlnLysThrAspCysAlaLeuArg 84
 399 GACGACCTGCTGCTGAGTGTCCACAGAAAGACAGATGCGCCCTCTGTG 448
 84 aLArgValValValHisLeuAlaValHisGlyHisThrAlaGluProGlu 100
 449 TCCGTGTGTGTCCACTGTGCGCTGCATGGGCGACCTGGCGAGCCCTGAA 498
 101 GluAlaGlyLysSerAspSerGluLeuGlnGlnSerArgAsnAlaSerLe 117
 499 GAAGCTGGAAGTCTGATTCAAGAACTCCAGAGGTCTAGACGCTCTCT 548
 117 uGlnAlaGlnValValLeuSerPheGlnAlaTyrProIleAlaArgCys 134
 549 CCAGGCCAGAGTGTGCTCTCTCCAGACCTACCCCATCGCCCGCTGTG 598
 134 ILeuLeuGluValGlnValProAlaAspLeuValGlnProGlyGlnSer 150
 599 CCTGTGTGAGGTCCAGGTGCCCTGACCTGTGTGACCTGTGTGCTGCTC 648
 151 ValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGluVal 167
 649 GTGGGTCTCTGCGTATTGACTGTTTCAGGCTAGCTTGGGGCTAGGT 698
 167 ILeuIleTyrSerTyrThrLysProArgTyrGlnLysGluLeuValLeu 184
 699 ACAGATCTGTGCTACAGAAAGCCAGGTACCAAGAAAGCTCAACTCA 748
 184 hTrGlnLeuProAspCysArgGlyLeuGlnValAlaArgAspSerIleGln 200
 749 CACAGCGCTGCTGCTGACGAGGGGTCTTGAAGTCCGGGACAGCACTCC 798
 201 SerCysTyrValLeuProTyrLeuAsnValSerThrAspGlyAspAsnVal 217
 799 AGCTGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 848
 217 ILeuLeuThrLeuAspValSerGluGlnGlnAspPheSerPheLeuLeu 234
 849 CTTCTGACACTGATGCTCTGAGGAGCAGAGCACTTACCTCTTACTGT 898
 234 YrLeuArgProValProAspAlaLeuLysSerLeuTyrTyrLysAsnLeu 250
 899 ACCTGCTGCAAGTCCCGATGCTCTCAAACTCTGTGTGACAAAACCTG 948
 251 ThrGlyProGlnAsnIleThrLeuAsnHisThrAspLeuValProCysLe 267
 949 ACTGGACCTCAGAACATTACTTTAAACACACAGACCTGGTCTCCCTGCT 998
 267 uCysIleGlnValTyrPheLeuGlnProAspSerGluArgValGluPheC 284
 999 CTGCATTCAGGTGTGTGTGCTAGAGCCAGACTCTAGAGGGTGCATTTCT 1048
 284 YsrPheAlaArgGluAspProGlyAlaHisArgAsnLeuThrHisIleAla 300
 1049 GCCCTTCCTCGGGAAGATCCCGGTGCACACAGGAACCTCTGGCACATAGCC 1098
 301 ArgLeuArgValLeuSerProGlyValTyrPheGlnLeuAspAlaProCys 317
 1099 AGGCTGCGGGTACTGTCCTCCAGGGGTATGGCAGATGCGCTTGTGCTG 1148
 317 sLeuProGlyLysValThrLeuCysTyrPheAlaProAspGlnSerProC 334
 1149 TCTGCGGGCAAGGTACACTGTGCTGCAGAGCACACAGACAGATCTCCT 1198
 334 YsgInProLeuValProProValProGlnLysAsnAlaThrValAsnGlu 350
 1199 GCCAGCCACTTGTGTCCACAGTGTCCCAAGAAAGCCACCTGTGATGTAG 1248
 351 ProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuCysValGlnVal 367

1249 CCACAGATTTCAGTTGTGTGGCAGGCCACCCCAACTCTGTGTCCAGT 1298
 367 IserThrTyrGluLysValGlnLeuGlnAlaCysSerTyrAlaAspSerL 384
 1299 GAGACCTGGAGAGAGTTTCACTGTCAAGGGTCTTGTGGCTGACTCTCT 1348
 384 euGlyProPheLysAspAspMetLeuLeuValGluMetLysThrGlyLeu 400
 1349 TGGGGCTTCAAGGATGATATGCTTGTAGTGTGAGATGAAACCGGCTC 1398
 401 AsnAsnThrSerValCysAlaLeuGluProSerGlyCysThrProLeuR 417
 1399 AACACACATCACTGTGTGCTTGGAAACCAAGTGTGTACACATCTCC 1448
 417 oSerMetAlaSerThrArgAlaAlaArgLeuGlyGluLeuLeuGlnAla 434
 1449 CAGCATGGCTCCACAGAGACTGTGCTGGGAGAGAGATGCTGTCAAG 1498
 434 sPheArgSerHisGlnCysMetGlnLeuTyrPasnAspAspAsnMetGly 450
 1499 ACTTCGATCACACACAGTGTATGACCTGTGGAACGATGACACATGGGA 1548
 451 SerLeuTyrAlaCysProMetAlaAspLysTyrIleHisArgTyrValLe 467
 1549 TCGCATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1598
 467 uValTyrLeuAlaCysLeuLeuLeuAlaAlaLeuPhePheLeuL 484
 1599 ACTATGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1648
 484 euleuLysLysAspArgArgLysAlaAlaAlaArgLysSerArgThrAlaLeu 500
 1649 TTTCTAAAAAGACCCGAGAAAGCCGCGCTGCTGCTGCTGCTGCTGCT 1698
 501 LeuLeuHisSerAlaAspGlyAlaGlyTyrGluArgLeuValGlyAlaLe 517
 1699 CTTCCTCACTCCGCGCAGGAGGGGCTGACGAGGCTGTGGGAGGACT 1748
 517 uAlaSerAlaLeuSerGlnMetProLeuArgValAlaValAspLeuTyrPS 534
 1749 GCGCTGCGGCTTGAGCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1798
 534 eArgArgGluLeuSerAlaHisGlyAlaLeuAlaTyrPheHisIleGln 550
 1799 GCCGCGCAGCTGAGCGGCGCAGGAGCCTTAGCCTGTTCCACACAG 1848
 551 ArgArgArgIleLeuGlnGlnGlyValValIleLeuLeuPheSerPr 567
 1849 CGAGCCGTATCTGTGAGGAGGGTGTGTGTATCTTCTCTCTGCGCC 1898
 567 cAlaAlaValAlaGlnCysGlnGlnIleThrLeuGlnLeuGlnThrValGluP 584
 1899 CCGGCGCGTGGGCACTGTACAGATGCTGCTGCTGCTGCTGCTGCTGCT 1948
 584 rGlyProHisAspAlaLeuAlaAlaTyrLeuSerCysValLeuProAsp 600
 1949 CCGGCGCCCATAGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1998
 601 PheLeuGlnGlyArgAlaThrGlyArgTyrValGlyValTyrPheAspG 617
 1999 TTCTGTCAAGGCGGGGCGACCGCGCTACGTGCGGGTCTACTTGTGAGG 2048
 617 YLeuLeuHisProAspSerValProSerProPheArgValAlaProLeuP 634
 2049 GCTGTGTACACAGACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2098
 634 hSerLeuProThrGlnLeuProAlaPheLeuAspAlaLeuGlnGlyGly 650
 2099 TCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2148
 651 CysSerThrSerAlaGlyArgProAlaAspArgValGluArgValThrG 667

2149 TGTCTCCACTTCCGGGGGACCCCGGACCGGGGTGAGACAGTACCA 2198
 667 nAlaLeuArgSerAlaLeuAspSerCysThrSerSerSerGluAlaProG 684
 2199 GGCGCTGGGGTCCGGCCCTGGACAGCTGACTTACTGCTGGAGAGCCCGCAG 2248
 684 lYcysCysGluGluThrPaspLeuGlyProCysThrThrLeuGlu 698
 2249 GCTGCTGGAGGAGATGGACCTGGACCTGACCTACACTAGAA 2292
 seq_name: gb:BC004759
 seq_documentation_block:
 LOCUS BC004759 2015 bp mRNA linear ROD 12-JUL-2001
 DEFINITION Mus musculus, clone MGC:6973 IMAGE:3154616, mRNA, complete cds.
 ACCESSION BC004759
 VERSION BC004759.1 GI:13435809
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 TITLE 1 (bases 1 to 2015)
 AUTHORS Strausberg, R.
 JOURNAL Direct Submission
 Submitted (21-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/
 Contact: villalona@bcm.tmc.edu.
 Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Muzny, D.M., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: http://image.linl.gov
 Series: IRM Plate: 6 Row: C Column: 3.
 Location/Qualifiers
 1. 2015
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="MGC:6973 IMAGE:3154616"
 /tissue_type="Mammary tumor. Brcal-/L: MMTV-Cre model. 10
 months old, gross tissue."
 /clone_lib="NCI_CGAP_Man3"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 177. 1880
 /codon_start=1
 /product="Unknown (protein for MGC:6973)"
 /protein_id="AAH04759.1"
 /db_xref="GI:13435810"
 /translation="MPVSMFLSLAIGRPVYSLEIRMEPDTARSGLSCHLMDG
 DIVLIPGSIOSAGCPVTPRTRELYLRCKQDKCALCVRVYHLAVHGMHAEPEA
 GKDSLEQESRNASLACQVLSFQATPILARCALLEVDVVOFGSVSAVDFE
 ASIGAEQVMSYTKPRYQKELNLQQLPDCGLEVRDSIQSCVLPMLNVSTGDGVL
 LTLDSBEDQFSELYLRPVDAKSLMYKNLTGQONTLNTDLVPLCAICDYMSELP
 DSREVECPREDPGARNLWHLARLVLSQWQLDAPCCLPQVTLQWQAPDQSPC
 OPIVPPVPOKNATVNEPDDQIVAGHRLCYQVSTWEEYQLOACIMADSLGFPKDM
 LVEAKGLNNTSYCALEPSCGTLPSMASTRAARGELLDERSHQOQLNNDWG
 SLWACPMDKYIHRWVLYWLAQLLAALFFFLIKRMTSPSPGSAACYISCKAG
 RPAATSGSTGCTOTLCPPRSASPSRSSPCPRSCRISLMHCREAPLPRGRDPTGWN
 E"

BASE COUNT 379 a 616 c 595 g 425 t
 ORIGIN
 alignment_scores:
 Quality: 3153.50 Length: 698
 Ratio: 5.265 Gaps: 1
 Percent Similarity: 85.817 Percent Identity: 85.673
 alignment_block:
 US-09-899-471-5 x BC004759 ..
 Align seg 1/1 to: BC004759 from: 1 to: 2015
 1 MetProValSerThrPheLeuLeuSerLeuAlaLeuGlyAlaArgAspProVal 17
 177 ATGCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 226
 17 lValValSerLeuGluAlaArgLeuMetGluProGlnAspThrAlaArgCys 34
 227 GGTCTCTCTCTGAGAGACTGATGAGACCTCAGACACTGACCTGCTGCTGCT 276
 34 erLeuGlyLeuSerCysHisLeuThrPaspGlyAspValLeuCysLeuPro 50
 277 CTCTAGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 326
 51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeu 67
 327 GGAAGCTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 376
 67 nThrGluLeuValLeuArgCysProGlnLysThrAspCysAlaLeuArgV 84
 377 GAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426
 84 aAlaArgValValHisLeuAlaValHisGlyHisThrAlaGluProGlu 100
 427 TCCGTGTGTGTCTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 476
 101 GluAlaGlyLysSerAspSerGluLeuGlnLysSerAlaArgAspAlaSer 117
 477 GAAGCTGGAAGTCTGATTCAGAACTCAGAGAGTCTAGAGAGCTGCTGCT 526
 117 uGlnAlaGlnValValLeuSerPheGlnAlaLysProIleAlaArgCysA 134
 527 CCAGGCCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 576
 134 lAlaLeuGluValGlnValProAlaAspLeuValGlnProGlyLysSer 150
 577 CCTGTGTGAGGTCCAGGTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 626
 151 ValGlySerAlaValPheAspCysPheGlnAlaSerLeuGlyAlaGluVal 167
 627 GTGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
 167 lGlnIleThrSerTyrThrLysProArgTyrGlnLysGluLeuAsnLeu 184
 677 ACAGATCTGCTCTACACAGAGCCAGGTACCAAGAGCTCAACCTCA 726
 722 CACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 776
 201 SerCysTrpValLeuProTyrPheLeuAsnValSerThrAspGlyAspAsnVal 217
 777 AGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 826
 217 lLeuLeuThrLeuAspValSerGluGluGlnAspPheSerPheLeuLeu 234
 827 CTTTGTGACACTGATGCTCTGTGAGAGACAGACTTGTAGCTTCTTACTGT 876
 234 yLeuLeuArgProValProAspAlaLeuLysSerLeuTrpTyrLysAsnLeu 250
 877 ACCTGCTCCAGTCCCGGATGCTCTCAAAATCTTGTGTGTTACAAAACCTG 926


```

251 ThrGlyProGlnAsnIleThrLeuAsnHisThrAspLeuValProCysIle 267
    |||||
927 ACTGAGCTCGAGAACATTACTTTAAACACACAGACCTGGTCCCTGCT 976
267 uCysIleGlnValTrpSerLeuGluProAspSerGluArgValGluPheC 284
    |||||
977 CTGCATTCAAGGTGGTGGCTAGAGCCACACTGTGAGAGGTGGAATCT 1026
284 ySPProPheArgIleAspProGluAlaHisArgAsnLeuTrpHisIleAla 300
    |||||
1027 GCCCTTCCTCCGGGAAGATCCCGGTGCACACAGAACTCTGGCACATAGACC 1076
301 ArgLeuArgValLeuSerProGlyValTrpGlnLeuAspAlaProCysG 317
    |||||
1077 AGGCTGCGGGTACTGTCTCCACAGGGGTATGGCAGCTAGATGCCGCTGCTG 1126
317 sLeuProGlyIysValThrLeuCysTrpGlnAlaProAspGlnSerProC 334
    |||||
1127 TCTGCGGGGCAAGGTAAACACTGTGCTGGAGGACACACAGATCTCCT 1176
334 ySGlnProLeuValProProValProGlnLysAsnAlaThrValAsnGlu 350
    |||||
1177 GCCAGCCACTGTGCTGACACAGTGCACAGAAAGAACGCCACTGTGAATGAG 1226
351 ProGlnAspPheGlnLeuValAlaGlyHisProAsnLeuCysValGlnVa 367
    |||||
1227 CCACAAAGATTTCAGTTGGTGGTGGAGCCACCCCAACCTCTGTGTCCAGGT 1276
367 lSerThrTrpIleLysValGlnLeuGlnAlaCysSerTrpAlaAspSerI 384
    |||||
1277 GAGACCTCGGAGAAAGTTCACCTGCTCAACCGCTGTGTGGGCTGACTCCT 1326
384 euGlyProPheLysAspAspMetLeuLeuValGluMetLysThrGlyLeu 400
    |||||
1327 TGGGGCCCTTCACAGATGATATGCTGTAGTGAGATGAAAACGGGCTCT 1376
401 AsnAsnThrSerValCysAlaLeuGluProSerGlyCysThrProLeuP 417
    |||||
1377 AACACACATCATGCTGTGCTTGGAAACCCACTGGCTGTACACCACTGGCC 1426
417 oSerMetAlaSerThrArgAlaAlaArgLeuGlyGluGluLeuLeuGlnA 434
    |||||
1427 CAGCATGGCCCTCCACGAGAGCTGCTGGCTGGGAAGAGAGTTGCTCAAG 1476
434 sPheArgSerHisGlnCysMetGlnLeuTrpAsnAspAspAsnMetGly 450
    |||||
1477 ACTTCGATCATCACACAGTGTATGCAAGCTGTGGAACGATGACAACATGGGA 1526
451 SerLeuTrpAlaCysPrometAspLysTrpIleHisArgArgTrpValIle 467
    |||||
1527 TCGCTATGGGCTGCTCCCATGTGACAAAGTACATCCACAGCGCTGGTCTCT 1576
467 uValTrpLeuAlaCysLeuLeuLeuAlaAlaLeuPhePheLeuL 484
    |||||
1577 AGTATGGCTGGCTGCTGCTACTGTGGCTGGCGGCTTTCTTCTTCCTCC 1626
484 euLeuLysLysAspArgArgLysAlaAlaArgLysSerArgThrAlaLeu 500
    |||||
1627 TTCTAAAAA..... 1636
501 LeuLeuHisSerAlaAspGlyAlaGlyTrpGluArgLeuValGlyAlaLe 517
    .....
1636 ..... 1636
517 uAlaSerAlaLeuSerGlnMetProLeuArgValAlaValaAspLeuTrpS 534
    .....
1636 ..... 1636
534 eArgArgGluLeuSerAlaHisGlyAlaLeuAlaTrpPheHisIleGln 550
    .....
1636 ..... 1636

```

```

551 ArgArgArgIleLeuGlnGluGlyValValIleLeuLeuPheSerI 567
    .....
1636 ..... 1636
567 uAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnLeuGlnIleThrValGluP 584
    .....
1636 ..... 1636
584 roGlyProHisAspAlaLeuAlaAlaTrpLeuSerCysValLeuProasp 600
    |||||
1637 ..GGACCGCATGAGCGCCCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 1684
601 PheLeuGlnGlyArgAlaThrGlyArgTrpValGlyValIleTrpPheAspG 617
    |||||
1685 TTCCTCAAGGCGGGGCGGACCGGCGCTACGTCGGGGTCTACTTGACAGG 1734
617 yLeuLeuHisProAspSerValProSerProPheArgValAlaProLeuP 634
    |||||
1735 GCTGTGCACACCGACTGTGTGCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1784
634 hSerLeuProThrGlnLeuProAlaPheLeuAspAlaLeuGlnGlyGly 650
    |||||
1785 TCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1834
651 CysSerThrSerAlaGlyArgProAlaAspArgValGluArgValThrG 667
    |||||
1835 TCTCCACTTCCCGCGGGGCGACCGCGGACCGGGTGGAGACAGTACAGCA 1884
667 nAlaLeuArgSerAlaLeuAspSerCysThrSerSerSerGluAlaProG 684
    |||||
1885 GCGGCTGCGGTCCGCCCTGTCACAGCTGTACTTCTAGCTCGGAGAACCCCGAG 1934
684 lYcysCysGluGluTrpAspLeuGlyProCysThrThrLeuGlu 698
    |||||
1935 GCTGTGCACAGATGGAGACTGGAGCTGTGACTACACTAGAA 1978

seq_name: gb_pat:AX350978
seq_documentation_block:
LOCUS AX350978 2094 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 12 from Patent WO0190358.
ACCESSION AX350978
VERSION AX350978.1 GI:18616354
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (sites)
AUTHORS Gorman,D.M.
TITLE Mammalian receptor proteins: related reagents and methods
JOURNAL Patent: WO 0190358-A 12-29-NOV-2001;
SCHERING CORPORATION (US)
FEATURES
source location/qualifiers
1..2094
/organism="unidentified"
/db_xref="taxon:32644"
/note="rodent: surmised Mus musculus"
BASE COUNT 247 a 251 c 393 g 385 t 918 others
ORIGIN

alignment_scores:
Quality: 2756.00 Length: 698
Ratio: 4.984 Gaps: 0
Percent Similarity: 79.226 Percent Identity: 71.490

alignment_block:
US-09-899-471-5 x AX350978
Align seg 1/1 to: AX350978 from: 1 to: 2094

1 MetProValSerTrpPheLeuLeuSerLeuAlaLeuGlyArgAsnProVal 17
|||||
::: ||| ||| |||||

```



```

617 yLeuEnHsProAspSerValProSerProPheArgValAlaProLeuP 634
1851 NTNTNTNCAATCCNGAYMSNGTNCNCNCCNTTYMGNGCNCNTNT 1900
634 heSerLeuProThrGlnLeuProAlaPheLeuAspAlaLeuGlnGly 650
1901 TYMSNYTNCNCNCAATYTCNCNTTYTTCGATGCTTCNCAGGNGN 1950
651 CysSerThrSerAlaGlyArgProAlaAspArgValGluArgValThrG 667
1951 TGYSNACNMSNGCNGNCCNGAYMGNGTNGARBMNGTNCNCA 2000
667 nAlaLeuArgSerAlaLeuAspSerCysThrSerSerSerGlnAlaProG 684
2001 KGCNTTNGNMSNGCNYTNGAYMSNTGYACNMSNMSNGARCCNCNG 2050
684 lYcysCysGlnGluTrpAspLeuGlyProCysThrThrLeuGlu 698
2051 GNTGYTGARGARGTGGAYTTCNGCNCNTGYACNMCNCTNGAR 2094

seq_name: gb-pat:AX073993
seq_documentation_block:
LOCUS AX073993 2255 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 1 from Patent WO0104304.
ACCESSION AX073993
VERSION AX073993.1 GI:12710225
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2255)
AUTHORS Presnell,S.R., Burkhead,S.K. and Powder,S.L.
TITLE Human cytokine receptor
JOURNAL Patent: WO 0104304-A 1 18-JAN-2001;
ZymoGenetics, Inc. (US)
FEATURES
source 1..2255
location/qualifiers
1..2255 /organism="Homo sapiens"
/db_xref="taxon:9606"
154..2232 /note="unnamed protein product"
/codon_start=1
/protein_id="CAC28435.1"
/translation="MPVWFLLSLALGRSPVLSIERLYGPODATHCSFGLSRLMDS
DIICLPEDIVAPGPVLAFTHLQTELVRCOKETDCDCLRAVHLAVHGMREPEDE
EKRGADSGVEEERNASIOAOVYLSPOAYTPARCVILEVOYPAALVQCGSYVY
DCEPAALGSERYRISYTPRYEKELNHTQQLPALPMNVNSAGDNVHLVNSSEOHF
GLSLTYNOVGPPKPRMHNLTGPOITLNTHTDVPCLCTIOVPLEPDSVTRNICPR
EDPAHONLWQAAELRLTLQSWLADACSLPAEALCMRARGDPCQPLVPELSPEN
VTVDKVEFLPLKGNPCVYNVNSSEKLOQLQEWADSGPLKDDVLLETREPORNR
SLALBESGCTSLPSKASTRAARGLEYLQDLOSGLQMDDDGALMACPMDKYTH
KRNALVYLACILFAALSLIILKKDHAKAARGRAVILITLSADSGSPRIYALASA
LCOLPLAVADIMSRRELSNCGPYANFPAQNRQTLQEGVYVLLSPGVALCSEMLQ
DGYSGPARGHDPFRASLSCVLPDFELQGRAPGSVGCDFDLHLHPDAVPALEFRTVPY
FTLPSPDLDFGALQDPAPRSRGRQERAEVSRALQPALDSYFHPGTPAPRGVGP
GAGPGAGDGT"
CDS
154..2232
/translation="MPVWFLLSLALGRSPVLSIERLYGPODATHCSFGLSRLMDS
DIICLPEDIVAPGPVLAFTHLQTELVRCOKETDCDCLRAVHLAVHGMREPEDE
EKRGADSGVEEERNASIOAOVYLSPOAYTPARCVILEVOYPAALVQCGSYVY
DCEPAALGSERYRISYTPRYEKELNHTQQLPALPMNVNSAGDNVHLVNSSEOHF
GLSLTYNOVGPPKPRMHNLTGPOITLNTHTDVPCLCTIOVPLEPDSVTRNICPR
EDPAHONLWQAAELRLTLQSWLADACSLPAEALCMRARGDPCQPLVPELSPEN
VTVDKVEFLPLKGNPCVYNVNSSEKLOQLQEWADSGPLKDDVLLETREPORNR
SLALBESGCTSLPSKASTRAARGLEYLQDLOSGLQMDDDGALMACPMDKYTH
KRNALVYLACILFAALSLIILKKDHAKAARGRAVILITLSADSGSPRIYALASA
LCOLPLAVADIMSRRELSNCGPYANFPAQNRQTLQEGVYVLLSPGVALCSEMLQ
DGYSGPARGHDPFRASLSCVLPDFELQGRAPGSVGCDFDLHLHPDAVPALEFRTVPY
FTLPSPDLDFGALQDPAPRSRGRQERAEVSRALQPALDSYFHPGTPAPRGVGP
GAGPGAGDGT"
BASE COUNT 369 a 748 c 703 g 435 t
ORIGIN
Alignment_scores:
Quality: 2389.50 Length: 693
Ratio: 4.148 Gaps: 5
Percent Similarity: 83.117 Percent Identity: 66.955
Alignment block:
US-09-899-471-5 x AX073993
Align seg 1/1 to: AX073993 from: 1 to: 2255

```

```

1 MetProValSerTrpPheLeuLeuSerLeuAlaLeuGlyArgAsnProVal 17
154 ATCCCTGCGCCCTGCTCTTCCTGCTTCCTGCTTCGACATCGGCGGAAGCCACT 203
17 lValValSerLeuGlnArgLeuMetGluProGlnAspThrAlaArgCys 34
204 GGTCTCTTCTCTGAGAGAGCTTGTGGCCCTCAGAGACGCTTACCTACTGCT 253
34 erLeuGlyLeuSerCysHisLeuTrpAspGlyAspValLeuCysLeuPro 50
254 CTCGGGCGCTCTCTGCGCCCTCTGGAGACATGGACATCTCGCTGCT 303
51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuG 67
304 GGGGACATCGCTGCTGCTCGCCGCCCTGCTGCGGCGCTTACCTACTGCA 353
67 nThrGlnLeuValLeuArgCysProGlnTrpAspCysAlaLeuArg 84
354 GACAGAGCTGCTGCTGAGTCCAGAGAGACCCAGCTGTGACCTCTGTC 403
84 alArgValValValHisLeuAlaValHisGlyHisTrpAlaGluProGlu 100
404 TCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 453
101 Glu.....AlaGlyLysSerAspSerGluLeuGlnGlnSerAr 113
454 GATGAGGAGAAAGTTTGGAGAGAGAGCTGACATGAGGAGGAGAGAGCTAG 503
113 gAsnAlaSerLeuGlnAlaGlnValValLeuSerPheGlnAlaTrpPro 130
504 GAATGCTCTCTCCAGGCGCAAGTGTGCTCTCTCTCCAGGCGCTTACCTTA 553
130 lAlaArgCysAlaLeuLeuGlnValGlnValProAlaAspLeuValGln 146
554 CTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 603
147 ProGlyLysSerValGlySerAlaValPheAspCysPheGlnAlaSer 163
604 TTGTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 653
163 uGlyAlaGlnValGlnLeuTrpSerTrpThrLysProArgTrpGlnLysG 180
654 AGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 703
180 lLeuLeuLeuThrGlnLeuProAspCysArgGlyLeuGlnValArg 196
704 AACTGACACACACACAGAGCTGCT..... 729
197 AspSerTrpLeuSerCysTrpValAlaLeuProTrpLeuAsnValSerThr 213
730 .....GCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 758
213 pGlyAspAsnValLeuLeuTrpLeuAspValSerGlnGlnAsnAspPhe 230
759 TGTGTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 808
230 erPheLeuLeuTrpLeuArgProValProAspAlaLeuLysSerLeuTrp 246
809 GCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 858
247 TyrLysAsnLeuTrpGlyProGlnAsnLethrLeuAsnHisThrAspLe 263
859 CACAAAACCTGACTGACGACCGACATCATCTTGAACACACACAGACT 908
263 uValProCysLeuCysLleGlnValTrpSerLeuGlnProAspSerGlu 280
909 GGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 958
280 rValGluPheCysProPheArgGlnAspProGlyAlaHisArgAsnLeu 296
959 GGACGAACATCTGCGCTTCAAGGAGAGACCCCGGACACACAGAACCTC 1008
297 TrpHisLleAlaArgLeuArgValLeuSerProGlyValTrpGlnLeuAs 313

```



```

1138 ACTGTGACAAAGTTCTTCGAGTTCCCAATGCTGTAAGGCCACCCCAACT 1287
363 UCYSValGlnValSerThrTrpGluLysValGlnLeuGlnAlaCysSer 380
1288 CTGTGTTACAGTGAACAGCTCGAGAAAGCTCAGCTGACAGAGTCTTCT 1337
380 rPalAspSerLeuGlyProPheLysAspMetLeuValGlnMet 396
1338 GGGCTAGCTCCCTGGGGCTCTCAAGAGATGCTCTGTTGGAGACA 1387
397 LysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluProSerGlyC 413
1388 CGAGGCCCCCAAGACAAACAGATCCCTCTGTGCTTGGAAACCAAGTGG 1437
413 sThrProLeuProSerMetAlaSerThrArgAlaAlaArgLeuGlyGlu 430
1438 TACTTACTACCCAGCAAGAGCTCCAGAGGCAAGCTGCGCTGGAGAGT 1487
430 LuleuLeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsn 446
1488 ACTTACTACAAAGACTGACAGCCAGCCAGCTGTCTGCACCTATGG...GAC 1534
447 AspAsnMetGlySerLeuTrpAlaCysProMetAspLysTrpIleHisAr 463
1535 GATGACTTGGAGAGCGCTATAGGCGCTGCCCATGAGCAAAATCATCCACA 1584
463 gArgTrpValLeuValTrpLeuAlaCysLeuLeuLeuAlaAlaLeu 480
1585 GCGCTGGGCGCTGCTGGTGGCTGCGCTGCTACTCTTTCGCGTGGCTT 1634
480 hePhePheLeuLeuLeuLysAspArgArgLys..... 491
1635 CCTCATCTCTTCTTCMAAAGATCACCGCAAGAGGTGGCTGAGGCTC 1684
492 .....AlaAlaArgGlySerArgThrAl 499
1685 TTGAAACAGACGCTCCGCTGGGGCGGCGCCAGGGGC...CGGGCGCG 1731
499 aleuLeuLeuHisSerAlaAspGlyAlaGlyTYrGluArgLeuValGly 516
1732 TCGGCTCCCTACTACAGCGATGACTGGGTTTGAGAGCGCTGGGGCG 1781
516 lAleuAlaSerAlaLeuSerGlnMetProLeuArgValAlaValAspLeu 532
1782 CCTGGCTGCGGCGCTGTGCTGACCTGCGCGCTGCGCTGAGACCTG 1831
533 TrpSerArgArgGluLeuSerAlaHisGlyAlaLeuAlaTrpPheHisH 549
1832 TGAAGCGTGTGTAAGTGAAGCGGAGGCGCGCTGGCTGTGTTCAAGC 1881
549 sGlnArgArgArgIleLeuGlnGluGlyValValIleLeuLeuPheS 566
1882 GCAGCGGCGCCAGACCTGTCAGAGAGGCGCGCTGTGCTTCTTCT 1931
566 erProAlaAlaValAlaGlnCysGlnIleTrpLeuGlnLeuThrVal 582
1932 CTCCGCGTGGCGGTGGCTGTGTCAGAGTGGTACAGAGTGGGTGCC 1981
583 GluPro.....GlyProHisAspAlaLeuAlaAlaTrpLeuSer 596
1982 GGGCGCGGGGCGGACGCGCGGACAGCGCTTCCGCGCTGCTCAGCTG 2031
596 sValLeuProAspPheLeuGlnGlnIleArgAlaThrGlyArgValGly 613
2032 CGTGCGCGCCGACTTCTTGAGGGCGGGGCGCGCGGAGCTACCTGGGG 2081
613 alTrpPheAspGlyLeuLeuHisProAspSerValProSerProPheArg 629
2082 CCTGCTTCGACAGGCTGTCCACCCGAGACCCGATACCCGCTTTCGCG 2131
630 ValAlaProLeuPheSerLeuProThrGlnLeuProAlaPheLeuAspAl 646
2132 ACCGTGCCGCTTCACAGTCCCTCCCAACTGCCAAGTTCCTGGGGGC 2181

```


1388 GGGCTGACCTCCGCGGCTCTCAAGAGATGCTGCTGAGACA 1387
 397 LysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluProSerGlyCys 413
 1388 CGAGGCCCCGAGGACACAGATCCCTGTGCTGCTGGAACCCAGTGGCTG 1437
 413 srhProLeuProSerMetAlaSerThrArgAlaAlaArgLeuGlyGluG 430
 1438 TACTTACTACTACCCAGCAAGACCTCCACAGGAGGAGCTCCCTTGGAGAGT 1487
 430 LuLeuLeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsnAsp 446
 1488 ACTTACTACCAAGACCTGAGTACAGGAGCTGTCTGACGCTATGG...GAC 1534
 447 AspAsnMetGlySerLeuTrpAlaCysProMetAspLysTyrIleHisArg 463
 1535 GATGACTGTGGAGCGCTATGAGCCCTGCCCATGAGCAATACATCCACAA 1584
 463 GATGTrpValLeuValTrpLeuAlaCysLeuLeuAlaAlaLeuP 480
 1585 GCGCTGGGCGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1634
 480 hePhePheLeuLeuLeuLysAspArgArgLys..... 491
 1635 CCTCATCTCTCTCTCAAAAGATCAAGCGAAGGGTGGCTGAGGCTC 1684
 492AlaAlaArgGlySerArgThrAl 499
 1685 TTGAAACAGACGTCCGCTCGGGGGGCGCCGACAGGGGC...CGCGCGGC 1731
 499 AlLeuLeuHisSerAlaAspGlyAlaGlyTyrGluArgLeuValGly 516
 1732 TCTGCTCTCTCTACTACGCGAGTACTCGGGTTTCAAGCCCTGCTGGGCG 1781
 516 AlaLeuAlaSerAlaLeuSerGlnMetProLeuArgValAlaValaLeuP 532
 1782 CCTGGGCTGGGCGCTGTGCGCAGCTGCGCTGCTGCTGCTGCTGCTGCTGCT 1831
 533 TrpSerArgArgGluLeuSerAlaHisGlyAlaLeuAlaTrpPheHisSH 549
 1832 TGGAGCGCTGTGAATGAGCGCGCAGGGGCGCTGCTGCTGCTGCTGCTGCT 1881
 549 sglArgArgArgIleLeuGlnGlyGlyValaValaIleLeuLeuPheS 566
 1882 GCGAGCGCGCGCAGACCCCTGCGAGGCGCGCTGCTGCTGCTGCTGCTGCT 1931
 566 expProAlaAlaValaAlaGlnCysGlnGlnTrpLeuGlnLeuGlnTrpVal 582
 1932 CTCCTGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1981
 583 GluPro.....GlyProHisAspAlaLeuAlaAlaTrpLeuSerCys 596
 1982 GGGCCCCGCGGCGCAGGCGCCGCGCAGGCGCTTCCGCGCTGCTGCTGCTG 2031
 596 sValLeuProAspPheLeuGlnGlyArgAlaThrGlyArgTyrValGly 613
 2032 CGTGTCTCCGAGCTTCTTGCAGGCGCGGCGCGCGCTGCTGCTGCTGCTGCT 2081
 613 AlTrpPheAspGlyLeuLeuHisProAspSerValProSerProPheArg 629
 2082 CTTGCTTGCAGAGGCTGCTCCACCGCGCGCGCTGCTGCTGCTGCTGCTGCT 2131
 630 ValAlaProLeuPheSerLeuProThrGlnLeuProAlaPheLeuAspAl 646
 2132 ACCGTGCGCTTCTTCACTGCTCCCTCCACAGTGCAGACTTCTGCGGCGC 2181
 646 AlLeuGlnGlyCysSerThrSerAlaGlyArgProAlaAspArgValG 663
 2182 CTCGACGACACCTCGCGCGCTGCTGCGGCGCGCTGCTGCTGCTGCTGCTGCT 2231
 663 LuArgValTrpGlnAlaLeuArgSerAlaLeuAspSerCys.....Thr 677
 2232 AGCAAGTGTCCGCGCGCTTCAAGCAGCGCTGAGATGCTGCTGCTGCTGCTGCT 2281

678 SerSerSerGluAlaProGly 684
 2282 CCGGGGACTCCCGCGCGGGA 2302
 seq_name: gb_pat:AX350973

seq_documentation_block: 2308 bp DNA linear PAT 06-FEB-2002
 LOCUS AX350973
 DEFINITION Sequence 7 from Patent WO0190358.
 ACCESSION AX350973
 VERSION AX350973.1 GI:18616349
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (sites)
 AUTHORS Gorman,D.M.
 TITLE Mammalian receptor proteins: related reagents and methods
 JOURNAL Patent: WO 0190358-A 7 29-NOV-2001;
 SCHERING CORPORATION (US)
 FEATURES
 source location/Qualifiers
 1..2308

CDS
 /organism="unidentified"
 /db_xref="taxon:32644"
 /note="Primate; surmised Homo sapiens"
 181..2292
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD2835.1"
 /db_xref="GI:18616350"
 /translation="MPVWFLSLAIGRSQWILSLRLVGPQDATHCSPLGISRLMDS
 DILIPDILVAPGCVLAPTHLOTLELALRCOKETDDCLRAVAVLHVHMEPEDE
 EKFGADLGVPEPNSLQAOVYSFOYPPARCVILEVQVPAALVQFOSVSVY
 DCPFAAGSEVRINSYTOPREKELNTHQQLDRCRLTEVNSIPSCRLPMVANSAG
 DNHVLVNSEQHFGSLRYNOVGPPPRRMKNLTGPQITLTHMDLVPCLTQVW
 PLEPDSVTRNICPFREDPPAHQNLQAAALRLITLQSMILDACSLPAEALCMRAG
 GDPCQPLVPLSMENVTVDVNSSEKIQLECLWADSLGPLKDDVLLLETGPODNRSL
 CALPEPGCTSLPESKASTRAARLGEYLIDQSGCQIQMDDDIGALWACPMQYIHR
 WALVWIACTLFPAAALSLILKLDHAKGRLIKODVRSQAARGRALLIYSADDSG
 FERVIGALASALCOLPLRVAVYDMSRELISAGCPVAFMFAQRQITLOEGVYVLLFSP
 GAVALLCSFMDGVSGPGAHGPHDAFASLSCLVLPFLQGRAGSYVAGCFDRLHPD
 AVPALFRTVPFTLPSQLPDLFGALQPPAPPSGRLOERAEVSRALQALDSYFHP
 GTSAPGRGVGAGGAGGAGT"
 241..2289

mat_peptide /product="unnamed"
 BASE COUNT 384 a 757 c 731 g 434 t 2 others
 ORIGIN

alignment_scores:
 Quality: 2365.00 Length: 707
 Ratio: 4.113 Gaps: 7
 Percent Similarity: 81.330 Percent Identity: 65.629

alignment_block:
 US-09-899-471-5 x AX350973

Align seg 1/1 to: AX350973 from: 1 to: 2308

1 MetProValSerTrpPheLeuLeuSerLeuAlaLeuGluAsnProVal 17
 181 ATGCTGTGCTCCCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 230
 17 ValAlaSerLeuGluArgLeuMetGluProGlnAspThrAlaArgCys 34
 231 GATCCTTCTCTGAGAGGCTGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 280
 34 erLeuGlyLeuSerCysHisLeuTrpAspGlyAspValLeuCysLeuPro 50
 281 CTCGGGCGCTTCCCTGCGCGCTGCTGCGGAGCTGACATACCTGCTGCTGCT 330
 51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuG 67


```

331 GGGGACATGCTGCTCGGGGCCCGCTGCGGCGCTACGACCTGCA 380
67 hThGluLeuValLeuArgCysProGlnIleThrAspCysAlaLeuArg 84
381 GACAGAGCTGGCTGAGGTGCGAGAGAGACCGGACTGTGACCTGTGTC 430
84 aAlaGValValAlaHisLeuAlaValHisGlyHisTrpAlaGluProGlu 100
431 TGGCTGTGGCTGTCCACTTGGCCGTGATGGCAGGAGAGAGCTGAA 480
101 Glu.....AlaGlySerSerAspSerGluLeuGlnIleSer 113
481 GATGAGGAAAGTTTGGAGAGACAGCTGACTTACGGGTGGAGAGCGCTAG 530
113 gAsnAlaSerLeuGlnAlaGlnValAlaLeuSerPheGlnAlaTyrPro 130
531 GAAATGCTCTCCAGGCCCAAGTGTCTCTCCAGGCTACCTTA 580
130 leaAlaArgCysAlaLeuLeuGlnValGlnValProAlaAspLeuValGln 146
581 CTGCGCCGCTGCTCTGCTGAGGTGCAAGTGTGCTGCTGCTGTGTGAG 630
147 ProGlnIleSerValGlySerAlaValPheAspCysPheGluAlaSerLe 163
631 TTGTGTAGTGTGTGGCTGTGTGATATGATGCTGTGAGGCTGCGCT 680
163 uGlyAlaGlnValGlnIleTrpSerTyrThrIleProArgTyrGlnIle 180
681 AGGAGTGTAGTGTGATGCTGTGTCTATATCTACAGCCAGGTACGAGAG 730
180 IuLeuAsnIleThrGlnIleLeuProAspCysArgGlyLeuGlnValArg 196
731 AACCAACACACACACAGCAGCTGCTGACTGCAAGGGGCTCGAAGTCTGG 780
197 AspSerIleGlnSerCysTyrValLeuProTrpLeuAsnValSerThrAs 213
781 AACGCAATCCCGAGCTGCTGGCCCTGCTGCTGCTGCTGCTGCTGCTG 830
213 pGlyAspAsnValLeuLeuThrLeuAspValSerGluGlnIleAspPhe 230
831 TGGTGTACACGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 880
230 ePheLeuLeuTyrLeuArgProValProAspAlaLeuIleSerLeuTyr 246
881 GCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 930
247 TyrIleAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAspLe 263
931 CACAAAACCTGACTGAGCGCAGATCATTTACTTGAACACACACACACT 980
263 uValProCysLeuCysIleGlnValTrpSerLeuGluProAspSerGlu 280
981 GGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1030
280 rValGluPheCysProPheArgGluAspProGluAlaHisAspLeu 296
1031 GGAGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
297 TrpHisIleAlaArgLeuArgValLeuSerProGluValTrpGlnLeuAs 313
1081 TGGCAACGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1130
313 pAlaProCysCysLeuProGlyIleValThrLeuCysTrpGlnAlaPro 330
1131 CGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1180
330 spGlnIleProCysGlnProLeuValProProValProGlnIleAsnAla 346
1181 GTGGGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1230
347 ThrValAsnGluProGlnAspPheGlnIleValAlaGlnHisProAsnLe 363
1231 ACNTGGAC..... 1239
363 uCysValGlnValSerThrTrpGluIleValGlnLeuGlnAlaCysSer 380
1240 .....GTGAACAGCTCGGAGAGCTCAGCTCAGCTGAGAGCTTGT 1279
380 rPalAspSerLeuGlyProPheIleAspAspMetLeuValGlnMet 396
1280 GGGCTGACTCCTCGGGCTCTCTCAAGAGAGTGTCTGTGTGAGAGACA 1329
397 LysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluProSerGlyC 413
1330 CGAGGCCCGCCAGAGCAACAGATCCCTGTGTCTGCTGGAACCCAGTGGCT 1379
413 sThrProLeuProSerMetAlaSerThrArgAlaAlaArgLeuGlyGlu 430
1380 TACTTACTACCCAGCAAGAGCTCCAGAGGGGAGCTGCTTGGAGAGT 1429
430 IuLeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsnAsp 446
1430 ACTTACTACAGAGCTCAGTACAGGCCAGTGTCTGCACTATAG...GAC 1476
447 AspAsnMetGlySerLeuTrpAlaCysProMetAspLysTyrIleHisAr 463
1477 GATGACTGTGGAGCGCTATGGGCTGCGCCCATGACAAATACATCATCACA 1526
463 gArgTrpValLeuValTrpLeuAlaCysLeuLeuLeuAlaAlaLeu 480
1527 GGGCTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1576
480 hPhePheLeuLeuLeuIleIleAspArgIle..... 491
1577 CCTGATCTCTCTCTCAAAAGATCCGCAAGAGGTGGCTGAGGCTC 1626
492 .....AlaAlaArgGlySerArgThrAl 499
1627 TTGAACAGAGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1673
499 aLeuLeuLeuHisSerAlaAspGlyAlaGlyTyrGluArgLeuValGly 516
1674 TCGCTCTCTTACTACGCGATGACTGCGGTGTGTGTGTGTGTGTGTGT 1723
516 IuLeuAlaSerAlaLeuSerGlnMetProLeuArgValAlaValAspLeu 532
1724 CCTGGGCTGCGGCTGTGCGAGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1773
533 TrpSerArgArgIleLeuSerAlaHisGlyAlaLeuAlaTrpPheHisI 549
1774 TGGAGCGGTGTGACTGAGCGGCGGCGGCGGCGGCGGCTGTGTGTGTGT 1823
549 sGlnArgArgArgIleLeuGlnGlnGlyValValIleLeuLeuPhe 566
1824 GCAAGCGGCGCAGACCTCGAGAGGCGGCGGCTGTGTGTGTGTGTGTGT 1873
566 eProAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnIleThrVal 582
1874 GTCCGCGGTGCGGCTGCTGCGGAGTGTGTGTGTGTGTGTGTGTGTGT 1923
583 GluPro.....GlyProHisAspAlaLeuAlaAlaTrpLeuSerC 596
1924 GGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGTGTGTGTGT 1973
596 sValLeuProAspPheLeuGlnGlyArgAlaThrGlyArgTyrValGly 613
1974 CGTGGCGCGGACTTGTTCAGGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2023
613 aLysPheAspGlyLeuLeuHisProAspSerValProSerProPheArg 629
2024 CCGTGTTCAGAGCTGTCTCACCGGAGCGGCGGCGGCGGCGGCGGCGGCG 2073
630 ValAlaProLeuPheSerLeuProThrGlnLeuProAlaPheLeuAspAl 646
2074 ACCGTGCGCGTTCACAGTGTGCTCCCACTTCCCACTTCCCACTTCCCACT 2123

```


213 pgiAspAsnValLeuLeuThrLeuAspValSerGluGluAspPhe 230
 803 TGGTGACAAAGTCATCTGTTGTAATGTCTCTAGAGACACACTTGC 852
 230 ePheLeuLeuThrLeuArgProValProAspAlaLeuLysSerLeu 246
 853 GCGTCTCTCTACTGGAATCAGGTCACAGGGCCCCCAAAACCCCGTGG 902
 247 TyrLysAsnLeuThrGlyProGluAsnThrLeuAsnHisThrAsp 263
 903 CACAAAACCTGACTGAGCCGACGATCATTTACCTTGAAACACACACACT 952
 263 uValProCysLeuCysIleGluValThrSerLeuGluProAspSerGlu 280
 953 GGTTCCTGCTGCTGTAATGAGGTGTGCTGTGAACTGACTGCTGTA 1002
 280 rGValGluPheCysProPheArgGluAspProGlyAlaHisArgAsnLeu 296
 1003 GGACGAACATCTGCTTCAAGGAGGACCCCGGACACACCAACATCTC 1052
 297 TrpHisIleAlaArgLeuArgValLeuSerProGlyValTrpGluLeu 313
 1053 TGGCAACCCCGCCACTGCGACTGCTGACCTGCAGAGCTGGCTGCTGA 1102
 313 pAlaProCysCysLeuProGlyLysValThrLeuCysTrpGluAlaPro 330
 1103 CGACACCTGCTGCTGCTGCGGACGAGGACGACTGCTGCGGCTGCTGG 1152
 330 spGlnSerProCysGlnProLeuValProProValProGluLysAsnAla 346
 1153 GTGGGGACCTCCAGCAGCAGCTGGTCCACCTCTCTGCGGAAAGCTC 1202
 347 ThrValAsnGluProGluAspPheGluLeuValAlaGlyHisProAsnLe 363
 1203 ACTGTGACAAAGTTCTGAGTTCCATTGCTGAAAGGCCACCTTAACCT 1252
 363 uCysValGluValSerThrTrpGluLysValGluLeuGluAlaCysSer 380
 1253 CTGTGTCAGGTGAACAGCTGCGAAGACCTGACGCTGCGAGCTGCTGT 1302
 380 rPalaAspSerLeuGlyProPheLysAspAspMetLeuLeuValGluMet 396
 1303 GGGCTGACTCTCTGGGCTCTCAAGAGATGCTCTACTGTTGAGACA 1352
 397 LysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluProSerGlyC 413
 1353 CGAGGCCCCAGGACACAGATCCCTGTGCTTGGAGACCCAGTGGCTG 1402
 413 sThrProLeuProSerMetAlaSerThrAlaAlaAlaArgLeuGlyGlu 430
 1403 TACTTTCATCCACGACAAAGCTCCACAGAGGACGCTGCTTGGAGAGT 1452
 430 LuLeuLeuGluAspPheArgSerHisGlnCysMetGlnLeuTrpAsnAsp 446
 1453 ACTTACTACAAAGACTGACAGTCCAGGCTGTCTGACGCTATGG...GAC 1499
 447 AspAsnMetGlySerLeuTrpAlaCysProMetAspLysTrpIleHisAr 463
 1500 GATGACTTGGAGGCTATGAGGCTGCGCCCATGAGCAAAATACATCCACAA 1549
 463 gATGTTrpValLeuValTrpLeuAlaCysLeuLeuAlaAlaLeuP 480
 1550 GCGCTGGGCTCTGCTGTGCTGCTGCTTCTTCTGCTGCTGCTGCTT 1599
 480 hPhePheLeuLeuLeuLysLysAspArgArgLys... 491
 1600 CCTCATCTCTCTCTCAAAAAGATCAACGCAAAAGGTGGCTGAGGCTC 1649
 491 491
 1650 TTGAACAGAGAGCTCCGCTCGGGGGTGTAGTGGAGACAGCGCTGGCGG 1699
 491 491

1700 AGGGCCGCCCCGGGAGACGAGGCTGTGCAGCTCACTCTCTCCCTCC 1749
 492AlaAlaArgGlySerArgThrAlaLeu 501
 1750 CATCTCTTTCTCCGCGAGCGCGCCAGGAGGC...GCGCGGCTCTCT 1796
 501 uLeuHisSerAlaAspGlyAlaGlyTrpGluArgLeuValGlyAlaLeu 518
 1797 CCTTACTGACCGGATCATCTGGGTTTCAGAGCGCTGTGGGCGGCTGG 1846
 518 lAspAlaLeuSerGlnMetProLeuArgValAlaValAspLeuPsr 534
 1847 CGTGGGCGCTGTCACCTGCTGCGGCGGCTGAGCTGAGCTGAGAC 1896
 535 ArgArgGluLeuSerAlaHisGlyAlaLeuAlaTrpPheHisGlnAr 551
 1897 CGTGTGAACACTGAGCGGACAGGCGCGCTGCTGTGTTACAGCGCAGC 1946
 551 ArgArgGlyLeuGluGluGlyValValIleLeuLeuPheSerPro 568
 1947 GCGCGAGACCTTCAGAGAGGCGGCTGTGTCTGTCTCTCTCTCTCCG 1996
 568 lAlaValAlaGlnCysGlnGlnTrpLeuGluLeuGlnThrValGluPro 584
 1997 GTGGGCTGGCGCTGTGAGAGAGTGGCTACAGAGTGGGCTGCCGCGCC 2046
 585GlyProHisAspAlaLeuAlaAlaTrpLeuSerCysValle 598
 2047 GGGGCGCAGCGCGCCGACGACGCTCTCCGCGCTGCTGCTGCTGCTGCT 2096
 598 uProAspPheLeuGluGlyArgAlaThrGlyArgTrpValGlyAlaTrp 615
 2097 GCGGACTTCTTTCAGAGGCGGCGCGCGGCTGCTGCTGCTGCTGCTGCT 2146
 615 hAspGlyLeuLeuHisProAspSerValProSerProPheArgValAla 631
 2147 TCGACAGGCTGCTCCACCGCGAGCGCTTACCGGCTTCTTCCGACCGTG 2196
 632 ProLeuPheSerLeuProThrGlnLeuProAlaPheLeuAspAlaLeuG 648
 2197 CCGCTTTCACACTGCGCTCCCACTCCAGACTCTCTGCGGCGCGCTGCA 2246
 648 nGlyGlyCysSerThrSerAlaGlyArgProAlaAspArgValGluArg 665
 2247 GCAGGCTCGGCGCGCGGCTTCCGCGGCTTCAAGAGAGAGCGGAGCAAG 2296
 665 aThrGlnAlaLeuArgSerAlaLeuAspSerCys.....ThrsSer 679
 2297 TGTCCGCGGCGCTTACCGCCAGGCTGTGATGATCTTCACTTCCATCCCGG 2346
 680 SerGluAlaProGly 684
 2347 ACTCCGCGCGCGGGA 2361
 seq_name: gb_pat:AX136341
 seq_documentation_block: 2499 bp DNA Linear PAT 30-MAY-2001
 LOCUS AX136341
 DEFINITION Sequence 263 from Patent EP1067182
 ACCESSION AX136341
 VERSION AX136341.1 GI:14272747
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2499)
 AUTHORS Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and
 Hayashl,K.
 TITLE Secretory protein or membrane protein
 Patent: EP 1067182-A 263 10-JAN-2001;
 JOURNAL Helix Research Institute (JP)

FEATURES
 source Location/Qualifiers
 1..2499
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 CDS 208..1011
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC39812.1"
 /db_xref="GI:14272748"
 /translation="MPVPMFLSLALRSRPSVLSELRVGPQDATHCSFGLSCRLMDS
 DLICLDGDIYPAPGVPVLAIPHILQTEILARCCKEDICDLRYAVHLAVGHMEEPDE
 ERFEGADSGVEEPRNLSLOAVLSFOAYPTARCVLEVOVPALVDFGSGSVYV
 DEFEALGSEVRIMSTQPRYEKELNTHQOLPALPMLVASDGDVHLVLYANVSEBQH
 GLSYMNQVGGPRKPMHKNLVRRPPOVSHCHRMVQRQCHIREDP"
 BASE COUNT 418 a 821 c 764 g 496 t
 ORIGIN

Alignment scores:
 Quality: 2219.00 Length: 778
 Ratio: 3.977 Gaps: 10
 Percent Similarity: 71.722 Percent Identity: 57.841

Alignment block:
 US-09-899-471-5 x AX136341

Align seg 1/1 to: AX136341 from: 1 to: 2499

```

1 MetProValSerTrpPheLeuLeuSerLeuAlaLeuGlyArgAsnProVal 17
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
208 ATGCGCTGCGCCCTGCTCTGCTGCTGCTGCGCAGCGGCGCGAAGCCAG 257
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
17 LValValSerLeuGluTrpLeuMetGluProGlnAspThrAlaArgCys 34
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
258 GGGCTCTTCTGAGAGGCTTGCGGCGCTCGAGAGCCCTACCCACTGCT 307
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
34 erLeuGlyLeuSerCysHisLeuTrpAspGlyAspValLeuCysLeuPro 50
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
308 CTCGGGCGCTCTGCGCGCTCTGCGGACAGTGAATCTGCTGCGCT 357
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeu 67
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
358 GGGGACATCGCTGCTGCGGCGCGCTGCTGCGGCGCTACGCACTGCA 407
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
67 nThrGluLeuValLeuArgCysProGlnLysThrAspCysAlaLeuArg 84
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
408 GAGAGACTGCTGCTGAGTCCAGAGAGAGCCAGCTGACCTCTGTC 457
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
84 alArgValValValHisLeuAlaValHisGlyHisTrpAlaGluProGlu 100
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
458 TGGGTGCGGCTGCTGCACTGCGCGCTGCGGCACTGGAGAGACTGCA 507
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
101 Glu.....AlaGlyLysSerAspSerGluLeuGlnLysSer 113
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
508 GATGAGAGAAAGTTTGGAGAGACAGCTGACAGGCGGTGGAGAGACTG 557
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
113 gAsnAlaSerLeuGlnAlaGlnValValLeuSerPheGlnAlaTrpPro 130
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
558 GAATGCTCTCTCCAGGCCCAAGTGTGCTCTCTCTCCAGGCGCTACCTA 607
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
130 LeuAlaArgCysAlaLeuLeuGlnValGlnValProAlaAspLeuValGln 146
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
608 CTCGCCGCTGCTGCTGCTGAGAGTGCAGTGCCTGCTGCTGCTGCG 657
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
147 ProGlyLeuSerValGlySerAlaValPheAspCysPheGlnAlaSer 163
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
658 TTGTGTCAGCTGCTGCGCTGCTGCTATATACAGCTGCTGAGGCGCT 707
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
163 uGlyAlaGluValGlnLysTrpSerTrpThrLysProArgTrpGlnLysG 180
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
708 AGGAGAGTGAAGATGCTGCTTACTACAGCCAGGCTACGAGAGG 757
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
180 IuLeuAsnLeuThrGlnGlnLeuProAspCysArgGlyLeuGlnValArg 196
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

758 AACCTAACCAACACACAGAGCTGCT..... 783
197 AspSerIleGlnSerCysTrpValLeuProThrLeuAsnValSerThr 213
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
784 .....GCCCTGCGCGCTGCTCAACGCTTCAGCAGA 812
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
213 pGlyAspAsnValLeuLeuThrLeuAspValSerGluGlnAspPhe 230
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
813 TGCTGACAAAGCTCATCTGTTCTGAATGCTCTGAGAGAGCAGCACTTCG 862
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
230 erPheLeuLeuTrpLeuArgProValProAspAlaLeuLysSerLeuTrp 246
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
863 GCTCTCTCCCTGACTGGAATCAGTCCAGGCGCCCAAAACCCCGGTGG 912
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
247 TyrLysAsn..... 249
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
913 CACAAAACCTGGTGAAGGCTCCCTCCCTCCCAAGTCCATTCCTCACTGAG 962
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
249 ..... 249
963 GCCGATGCTGTGCAAGAGCAGTGCATATCAGAGAGATCTTGAA 1012
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
249 ..... 249
1013 GAGGACTCAACCCCAAGCAGGAAATGTGTGGGGAATTCTGCTTCC 1062
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
249 ..... 249
1063 TGGTTCCTTGACTTGGCTCTCCCTCTTCTCTTATCTTCCAAAC 1112
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
250 .....Leu..ThrGlyProGlnAsnLysThrLeuAsn 259
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1113 TCCTTCTTATTTGTTCCACAGACTGAGCGCAGATATTAACCTTGAAAC 1162
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
260 HisThrAspLeuValProCysLeuCysIleGlnValTrpSerLeuGluPr 276
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1163 CACAGAGCTTAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCAAC 1212
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
276 oAspSerGluArgValGluPheCysProPheArgGluAspProGlyAlaH 293
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1213 TGACTCCGTTAGACAGACATCTGCCCCCTTACGAGAGAGACCCCGCGAC 1262
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
293 IsArgAsnLeuTrpHisIleAlaArgLeuArgValLeuSerProGlyVal 309
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1263 ACCAGAACCTTGCAAGCGCCGAGCTGCGAGCTGACCTGACAGAC 1312
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
310 TrpGlnLeuAspAlaProCysCysLeuProGlyLysValThrLeuCysTr 326
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1313 TGGCTGCTGAGCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1362
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
326 pGlnAlaProAspGlnSerProCysGlnProLeuValProProValProG 343
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1363 GCGGCTCGCGGGGGGAGCCCTGCGGAGCCATGCTCCACCCGCTTCT 1412
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
343 LnyLysAsnAlaThrValAsnGluProGlnAspPheGlnLeuValAlaGly 359
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1413 GGGAGAACGTCAGTGGACAAAGTTCTGAGTTGATGCTGAAAGGC 1462
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
360 HisProAsnLeuCysValGlnValSerThrTrpGluLysValGlnLeuG 376
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1463 CACCTTAACCTGTGTTCAAGTGAACAGCTCGGAGAGCTGACGTGCA 1512
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
376 nAlaCysSerTrpAlaAspSerLeuGlyProPheLysAspAspMetLeu 393
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1513 GGAGCTCTTGTGGCTGCTCCCTGCGGCTCTCAAGAGAGATGTGCTAC 1562
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
393 euValGluMetLysThrGlyLeuAsnAsnThrSerValCysAlaLeuGlu 409
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1563 TGTGAGACAGACAGAGCCCGCAGACAAAGATCCCTGCTGCTTGGA 1612
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
410 ProSerGlyCysThrProLeuProSerMetAlaSerThrArgAlaAlaAr 426
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```



```

1613 CCCAGTGGCTTACTTCACTACCCAGCAAGCCTCCAGC..... 1651
426 gLeuGlyGluLeuLeuGlnAspPheArgSerHisGlnCysMetGlnL 443
1652 .....C 1652
443 euTPAsnAspAspAsnMetCylSerLeuTPAlaCysProMetAspLys 459
1653 TATGG...GACGATGACTTGGAGCGCTATGGCGCTGCCCATGACAA 1699
460 TyrLeuHisArgTyrValLeuValTyrPheAlaCysLeuLeuAl 476
1700 TACATCCACAAGCCGTGGCCCTGCTGGCGGCTGCTACTCTTGG 1749
476 aAlaAlaLeuPhePheLeuLeuLysLysAspArgLys.... 491
1750 CGGTGGCTTCCCTCATCCCTCTTCAAAAAGATCACGGCAAGGGT 1799
492 .....AlaAlaArgGly 495
1800 GGCTGAGGCTTGTAAACAGAGACGTCGCGGGCGCGCCAGGGCG 1849
496 SerArgThrAlaLeuLeuLeuHisSerAlaAspGlyAlaGlyTyrGluAr 512
1850 ...CGCGCGCTGCTGCTCTACTACAGCCGATGACTCGGGTTTCAGAGCG 1896
512 gLeuValGlyAlaLeuAlaSerAlaLeuSerGlnMetProLeuArgVal 529
1897 CCGGTGGGGCCCTGGCGGTGGCGCTGCCAGCTGCGCGGTGGCGGTG 1946
529 LaValAspLeuTyrSerArgArgGluLeuSerAlaHisGlyAlaLeuAla 545
1947 CCGTAGACGTGTGAGCGCGTGTGAACAGAGCGCCAGCGCCCGTGGCT 1996
546 TrpPheHisHisGlnArgArgArgGluLeuGlnGlyValVal 562
1997 TGGTTTCACCGCCAGCGCCAGACCTGTCAGAGAGCGCGGTGGT 2046
562 eLeuLeuPheSerProAlaAlaAlaGlnCysGlnGlnTyrPheGlnL 579
2047 CTTCGCTTCTCTCCCGGTGGCGGTGGCGGTGGCGGTGGCGGTG 2096
579 euGlnThrValGluPro.....GlyProHisAspAlaLeuAla 592
2097 ATGGGTGTCCGGGCGCGGGCGAGCGCCGCGACAGCCCTCCCGCGC 2146
593 TrpLeuSerCysValLeuProAspPheLeuGlnGlyArgAlaThrGlyAr 609
2147 TCGCTCAGCTGCGGTGCGCGGTCTTTCAGGGCGCGCGCGCGCAG 2196
609 gTyrValGlyValTyrPheAspGlyLeuLeuHisProAspSerValPro 626
2197 CTACGTGGGGGCGCTGCTTCAGACAGCTCTCCACCGGAGCGGTACCCG 2246
626 euProPheArgValAlaProLeuPheSerLeuProThrGlnLeuProAla 642
2247 CCTTTTCCGACCGGTGGCTTTCACACAGCCCTCCCACTGCCAGAC 2296
643 PheLeuAspAlaLeuGlnGlyCysSerThrSerAlaGlyArgProAl 659
2297 TTTCCTGGGGGCGCTGACAGCCTCGCGCGCGCGCTCCGCGCGCTCA 2346
659 aAspArgValGluArgValThrGlnAlaLeuArgSerAlaLeuAspSerC 676
2347 AGAGAGCGGAGAGAGTGTCCGCGGCTTTCAGCAGCCCTGATAGCT 2396
676 ys.....ThrSerSerGluAlaProGly 684
2397 ACTTCAATCCCGCGGAGACTCCGCGCGCGGCA 2428

```

seq_name: gb_pat:AX073996
seq_documentation_block:

```

LOCUS AX073996 1753 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 4 from Patent WO0104304.
ACCESSION AX073996
VERSION AX073996.1 GI:12710228
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1753)
AUTHORS Presnell,S.R., Burkhead,S.K. and Powder,S.L.
TITLE Human cytokine receptor
JOURNAL Patent: WO 0104304-A 4 18-JAN-2001;
ZymoGenetics, Inc. (US)
FEATURES
source Location/Qualifiers
1..1753
/organism="Homo sapiens"
/db_xref="taxon:9606"
<2..1729
/note="unnamed protein product"
/codon_start=1
/Protein_id="CAC28436.1"
/db_xref="GI:12710229"
/translation="EEPRNASLQAQVVLSPQAPTRACVLEVOVPALVQFGSVGS
VYVDFEALGSEVRIMSYTOPREKELNHTQOLPALPMLNSADGNVHLVNSE
OHFGSLYMNQVQPPRMHKNLNGPOLITHTDLPCLQVMPLEDSVNTIC
PFREDPRAHONIMOARLELTLQSLDLAPCSLPABALCWAPRGDPCQPLVPLS
WENYVDVNSSKERLQDQECLEMDSTLPLQDVLLLETRGPDNDNSLCALPESGCTSLP
SKASTRAARLSEKYLQDLOSGOCLQDMDLDDILAMCPMDKXYHKRMALVWLACILFA
AALSLILILKDKHAKGMLRLIKODVRSAAARALALILYSADSGFERLVGALASAL
COLPLRVAVDMSRRELNSAQGVAMVAHQAORRTLOEGVVLVLFSPGAVLCSMID
GVSGAGHPHAPRASISCVLPDFTQGARPSYVACPDRLIHPDAVPALEFRVYPR
TLPSQLPFLCALQOPRARPSGRLEAEQVSRALDPALDSTFHPETPAVRGVRG
AGPGAGDGT"
BASE COUNT 287 a 584 c 548 g 334 t
ORIGIN
alignment_scores:
Quality: 1888.50 Length: 594
Ratio: 4.053 Gaps: 7
Percent Similarity: 78.451 Percent Identity: 63.805
alignment_block:
US-09-899-471-5 x AX073996 ..
Align seg 1/1 to: AX073996 from: 1 to: 1753
110 GInGlnSerArgAsnAlaSerLeuGlnAlaGlnValValLeuSerPheG 126
2 GAGGAGCCTAGAGATGCTCTCTCCAGGCCCAAGTGGCTCTCTTCCA 51
126 nAlaTyrProIleAlaArgCysAlaLeuLeuGluValGlnAlaProAla 143
52 GGCGTACCTACTAGCCCGGTGCGTCTGTGAGAGTGAAGTGGCTGTG 101
143 sPLeuValGlnProGlyGlnSerValGlySerAlaValPheAspCysPhe 159
102 CCTCTGTGAGTTTGGTTCACACTGTGCGCTGTGCTATACAGCTGTC 151
160 GluAlaSerLeuGlnValGlnGlnIleTyrSerTyrThrLysProAr 176
152 GAGGTCGCCCTAGGAGTAGAGTACGATCTGCTCTATACTACAGCCAG 201
176 gTyrGlnLysGluLeuAsnLeuThrGlnGlnLeuProAspCysArgGlyL 193
202 CTACGAGAGAGAACTCAACACACAGCAGCTGCT..... 238
193 euGluValAlaArgAspSerIleGlnSerCysTyrValLeuProTyrLeuAsn 209
239 .....GCCCTGCGCTGCTCAAC 256
210 ValSerThrAspGlyAspAsnValLeuLeuThrLeuAspValSerGluG 226

```


[illegible]

ORIGIN

Alignment scores:

Quality: 1675.50 Length: 693
 Ratio: 3.757 Gaps: 5
 Percent Similarity: 64.358 Percent Identity: 47.186

Alignment block:

US-09-899-471-5 x AX073995 ..

Align seg 1/1 to: AX073995 from: 1 to: 2076

```

1 MetProValSerTriPheLeuLeuSerLeuAlaLeuGlyArgAsnProVal 17
1 ATGCCGNTCCNTGCTTYYTNTNMSNTYNGCNYYTNGMNMNSNCCNGT 50
17 ValValSerLeuGluArgLeuMetGluProGlnAspThrAlaArgCys 34
51 NGTNYTMSNTTNGARMGNTTNGTNGCNCARAGAYGCACNCAATGYW 100
34 LeuGluGlyLeuSerCysHisLeuTriPaspGlyAspValLeuCysLeuPro 50
101 SNCCNGNTYMTMSNTGYMTNTGGAATMSGATVHTYNTGYTNCN 150
51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuG1 67
151 GGNCAATHTGTCNCNCNCNCNCNCNTNTYNGCNCNACNCAATYTCA 200
67 nThGluLeuValLeuArgCysProGlnIuSerThrAspCysAlaLeuArgV 84
201 RACGAGATYNTNTNMGNTGTCARARARACNAGTGGAYTNTGYT 250
84 AlaArgValValHisLeuAlaValHisGlyHisTriPalaGluProGlu 100
251 TNGNGTNGCNCNTCAATYNTNGCNCNTGNCATYGGAGARARCCNGAR 300
101 Glu.....AlaGlyLysSerAspSerGluLeuGlnIuSerAr 113
301 GAYGARGARARARTTYGGNGCNCNCNCAGAYWSNGGNTNGARARCCNMG 350
113 GAsnAlaSerLeuGlnAlaGlnValValLeuSerPheGlnAlaArgProI 130
351 NAATGCMNSNTTNCARGCNCARGTGNTNTNMSNTTYTCARGCNTATCCNA 400
130 LeuAlaArgCysAlaLeuLeuGlnValGlnValProAlaAspLeuValGln 146
401 CNGCNGMTGYGTNTNTNGARGTNCARGTNCNCNCNCNTYNTNGTNCAR 450
147 ProGlyLysSerValGlySerAlaValPheAspCysPheGluAlaSerLe 163
451 TTYGNCARMSNGTNGCNCNCNTNTAYGATGYTGYGARCCNCNNTY 500
163 uGlyAlaGluValGlnIleTriPserTyrThrLysProArgTyrGluLysG 180
501 NGGMSNGARGTNMGNTATHTGMSNTATACNCAACCCNMGNTAAGARARG 550
180 IuLeuAsnLeuThrGlnIuLeuProAspCysArgGlyLeuGluValArg 196
551 ARYTNAAYCAVACNCAARARNTNCN..... 576
197 AspSerIleGlnSerCysTriPValLeuProTriPLeuAsnValSerThra 213
577 .....GCNTYNCNCNTGTYTNAAYGTNMSNCCNGA 605
213 pGlyAspAsnValLeuLeuThrLeuAspValSerGluGluGlnAspPhe 230
606 YGGAGATAAGTNCATYNTNGTNTNAAGTNNNSNAGARCARCAATTYG 655
230 eirPheLeuLeuTyrLeuArgProValProAspAlaLeuLysSerLeuTri 246
656 GNTYMTSNTYNTATYTGGAAYCARGTNCARGCNCNCNACARCCNMGNTG 705

```

```

247 TyrLysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAspLe 263
706 CAYARARAYTNAACNGCNCNCARATHATACNTYNTNAAYCAVACNAGAYT 755
263 uValProCysLeuCysIleGlnValTriPserLeuGluProAspSerGlu 280
756 NGTNCNTGYTNTGYTATHTCARGTNTGCCNTYNTNARCCNAGAYWSNGT 805
280 rValGluPheCysProPheArgGluAspProGlyAlaHisArgAsnLeu 296
806 GNCNNAATHTHTGCCNTTNTNGNARCAATCCNMGNCNCARARAYT 855
297 TrpHisIleAlaArgLeuArgValLeuSerProGlyValTriPLeuLeu 313
856 TGCCARCCNCCNMGNTYNTMGNTYNTNACNTNCARMSNTGTYTNTCA 905
313 PalaProCysCysLeuProGlyLysValThrLeuCysTriPAlaProA 330
906 YGCNCCNTGYMSNTYNTCCNCGNARGCNCNTNTGYTGMGNGCNCNG 955
330 spGlnSerProCysGlnProLeuValProProValProGlnIuLysAsnAla 346
956 GNGCNGATCCNTGYTCARCCNTNTGCCNCCNTYMTMSNTGGARARAYG 1005
347 ThrValAsnGluProGlnAspPheGlnLeuValAlaGlyHisProAsnLe 363
1006 ACNGTNGAYARAGTNTYNTNGARTTCYCNNTYNTNARGCNCAVCAAYT 1055
363 uCysValGlnValSerThrTriPLeuLysValGlnLeuGlnAlaCysSer 380
1056 NTGYTNCARGTNAAYWSNMSNAGARARAYTNCARTYTCARARTGYTNT 1105
380 rPalaAspSerLeuGlyProPheLysAspAspMetLeuLeuValIuMet 396
1106 GGCCNCAATYMSNTYNTNGCNCNTTNAARCAATGAYGTYTNTYNTNARAC 1155
397 LysThrGlyLeuAsnAsnThrSerValCysAlaLeuGlnIuProSerGlyC 413
1156 MGNGCNCNCARCAVAYAYMGMSNTYNTGYCNCNTNGARCCMWSMGNTG 1205
413 sThrProLeuProSerMetAlaSerThrArgAlaAlaArgLeuGlyGluG 430
1206 YACNMSNTYNTCCMNSNARCCNMSNACNMGNCNCNNGNTNGNGART 1255
430 IuLeuLeuGlnAspPheArgSerHisGlnCysMetGlnLeuTriPAsn 446
1256 ATYNTTNCARAGAYTNCARMSNCGNCARFTGYTNCARTYTTGG...CAT 1302
447 AspAspMetGlySerLeuTriPalaCysProMetAspLysTyrIleHisAr 463
1303 GAYGAYTNTGNCNCNTYNTGGCNCNTGYCNCNTGAYARATYATCAVAA 1352
463 gArgTriPValLeuValTriPLeuAlaCysLeuLeuAlaAlaLeuP 480
1353 RMGTGGCCNTTNGTNTGTYTNGCNTGYTNTYNTYNTYNTYNTYNTYNT 1402
480 hePhePheLeuLeuLeuLysLysAspArgArgLysAlaAlaArgLysSer 496
1403 SYTNTATHTNTYNTYNTAARARARAGAYCAAGCNAARGCNCNCNNGN 1452
497 ArgThrAlaLeuLeuLeuHisSerAlaAspGlyAlaGlyTyrGluArgLe 513
1453 MGNGCNCNTYNTYNTYNTAYWSNCGNAGAYMSNGNTYTGARGMNTY 1502
513 uValGlyAlaLeuAlaSerAlaLeuSerGlnMetProLeuArgValAla 530
1503 NGTNGCNCNTYNTGCMWSNCGNNTNTGYTCARTYTCNCNTYNTNGNCG 1552
530 AlaAspLeuTriPserArgArgGluLeuSerAlaHisGlyAlaLeuAlaTri 546
1553 TNGAYTNTGMSNMGNCARFTTMSNCGNCARGCNCNCNCGTNCNTGG 1602
547 pheHisIleGlnArgArgArgIleLeuGlnIuGlyValValIleLe 563

```



```

313 PALAProCysCysLeuProGlyLysValThrLeuCysTrpGlnAlaProA 330
|||||
951 YGCNCCNGYWNNTYTCNCCNGCNCARCGCNCNTYNTGTTGCGMNCNCNG 1000
330 spGlnSerProCysGlnProLeuValProProValProGlnLysAsnAla 346
|||||
1001 GNGCGATCCNTGTCARCCNTYNTGCCNCCNTYNTMSNTGGARAAAGTN 1050
347 ThrValAsnGluProGlnAspPheGlnLeuValAlaGlyHisProAsnLe 363
|||||
1051 ACNGTNGAY ..... 1059
363 uCysValGlnValSerThrTrpGluLysValGlnLeuGlnAlaCysSerT 380
|||||
1060 .....GTNAAYMSWMSNGARARARYTNCARATGCAAGATGYTNT 1099
380 rPALAAspSerLeuGlyProPheLysAspPheLeuLeuValGlnMet 396
|||||
1100 GGGCGATGWSNTYTCNCCNTYNTAARGATGATGTYNTYNTYNTGARRCN 1149
397 LysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluProSerGlyCys 413
|||||
1150 MGNCGNCCNCARGAAYAAAYMNSNTYNTGYCNYTNGARCCNMSNGNTG 1199
413 sThrProLeuProSerMetAlaSerThrArgAlaAlaArgLeuGlyGluG 430
|||||
1200 YACNMSNTYTCNCCNWSNARCGMNSNACNMNGCNCNMNGTYTNGCNGART 1249
430 LuLeuLeuGlnAspPheArgSerHisGlnCysMetGlnLeuTrpAsp 446
|||||
1250 AYTNTYTCARATYTCNARMSNGCARTGYTTCNARTNTG...GAY 1296
447 AspAsnMetGlySerLeuTrpAlaCysPrometAspLysTrpIleHisAr 463
|||||
1297 GAYGAYTNGCNCNTYNTGCCNTGYCCNATGATGAATATATATACAA 1346
463 gArgTrpAlaLeuValTrpLeuAlaCysLeuLeuAlaAlaLeu 480
|||||
1347 RMGNTGGCNCNTYNTGTYTNCNTGYTNTYNTYTCNCCNCCNTYNTW 1396
480 hepPheLeuLeuLeuLysLysAspArgLys ..... 491
1397 SNTYATHTYNTNTNARARAGATCAACNARCGNTGTYTNTMNTN 1446
492 .....AlaAlaArgLysSerArgThrAlaLe 500
1447 YTNARCARGAYGTNMGNSNGCNCNGCNCNMNGCNCNGCNCNTY 1496
500 uLeuLeuHisSerAlaAspGlyAlaGlyTrpGluArgLeuValAla 517
|||||
1497 NYTNTYNTAAYWNSNGCNGAYGAYWNSNGNTTYGARMCNTYNTGTCN 1546
517 euAlaSerAlaLeuSerGlnMetProLeuArgValAlaValAspLeuTrp 533
|||||
1547 TNGCWSNCCNTYNTGTCARNTNCCNTYNTMGNTGNCNGNTGAYTNTG 1596
534 SerArgArgGluLeuSerAlaHisGlyAlaLeuAlaTrpPheHisIsgl 550
|||||
1597 MSNMGNGMNGAYTNTMSNCCNARCGCNCNTGTCGTTTCAAGCNCA 1646
550 nArgArgArgIleLeuGlnGluGlyValValIleLeuLeuPheSerP 567
|||||
1647 RMGWSNCCNARACNTYNTNARCARGNGCNGTNGTYNTYNTTYSNC 1696
567 roAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnLeuGlnThrValGlu 583
|||||
1697 CNGCGNCCNTYNTGTCGNTGTCGNTGTCGNTGTCGNTGTCGNTGTCG 1746
584 Pro.....GlyProHisAspAlaLeuAlaAlaTrpLeuSerCysVa 597
|||||
1747 CCNCGNCCNCAATGNCNCNCAATGNCNCNCAATGNCNCNCAATGNCNC 1796
597 lLeuProAspPheLeuGlnGlyArgAlaThrGlyArgTyrValGlyVal 614

```

```

1797 NYTTCNCGAATYTTTCNARCGMNGCNCNCCNMSNTAYTNGNGCWT 1846
614 YrPheAspGlyLeuLeuHisProAspSerValProSerProPheArgVal 630
|||||
1847 GYTYGAYMGNNTYNTYNTCAACNGAYGCGNTGNCNCCNTNTTYMNAON 1896
631 AlaProLeuPheSerLeuProThrGlnLeuProAlaPheLeuAspAla 647
|||||
1897 GTNCCNCTNTYACNTYTCNCCNCAATTCNCCNCAATTTTNGNCNTT 1946
647 uGlnGlyGlyCysSerThrSerAlaGlyArgProAlaAspArgValGlu 664
|||||
1947 NCARCARCCNNGCNCNCCNMGNSNGMNGNTYNTNCARGARMGNCNGAR 1996
664 rGlyAlaThrGlnAlaLeuArgSerAlaLeuAspSerCys.....ThrSer 678
1997 ARGTNMSMNGCNCNTYNTNCARCCNCCNTYNTGAYMSNTAYTTCAYCCNCCN 2046
679 SerSerGluAlaProGly 684
2047 GGNACNMSNCCNCCN 2064

seq_name: gb_pat:AX073998
seq_documentation_block:
LOCUS AX073998 1725 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 6 from Patent WO0104304.
ACCESSION AX073998
VERSION AX073998.1 GI:12710230
KEYWORDS
SOURCE
ORGANISM synthetic construct.
artificial sequence.
REFERENCE
1 (bases 1 to 1725)
Presnell,S.R., Burkhead,S.K. and Powder,S.L.
Human cytokine receptor
Patent: WO 0104304-A 6 18-JAN-2001;
JOURNAL
ZymoGenetics, Inc. (US)
FEATURES
source
1..1725 Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="This degenerate sequence encodes the amino acid
sequence of SEQ ID NO:5."
misc-feature
1..1725
/note="n = A,T,C or G"
variation
1..1725
/note="N is any nucleotide."
BASE COUNT 192 a 226 c 332 g 216 t 759 others
ORIGIN
alignment_scores:
Quality: 1314.50 Length: 593
Ratio: 3.621 Gaps: 6
Percent Similarity: 61.214 Percent Identity: 45.025
alignment_block:
US-09-899-471-5 x AX073998
Align seg 1/1 to: AX073998 from: 1 to: 1725
110 GlnGlnSerArgAsnAlaSerLeuGlnAlaGlnValValLeuSerPheG 126
|||||
1 GARGARCCNNGNNAAYGNCNWSNTYNTNCARCCNCAAGTNGTYTMSNTTCA 50
126 nAlaIyrProIleAlaArgCysAlaLeuLeuGluValGlnValProAla 143
|||||
51 RGCNTAACCNAACNCGNMGNTGYGTYNTYNTGARGTNCARGTNCNCCNG 100
143 splLeuValGlnProGlnSerValGlySerAlaValAlaPheAspCysPhe 159
|||||

```


460 TTTTLeHsArGaTTPvAlLeuValTPrLeuAlAcYSLeuLeuAl 476
|||||..... ||||| |||||
952 TAYTThCAyARmGNTGGcCYTNGTNGTcGNTGYYTNTNTTGYcC 1001
476 aAlaAlaLeuPheHheLeuLeuLeuLySArGArGdYs... 491
||||| ||||| |||||
1002 NGCGcCYTmMSNTYNTATHTYNTYNTYNTAAARGAACAYcCNAARgGNT 1051
492AlaAlaArgLySArG 496
1052 GGYTmGNTYNTAAARCArGATYTNmGmNSNGcCGcCGcCMGNGN 1101
497 ArgThAlaLeuLeuLeuHISerAlaSpLyAlaGlyTyrgLuarGle 513
||||| ||||| |||||
1102 mGNCcCNCNTYNTYNTATAYmSGNCNGAYGAYmSGNGNTTYGARmGNT 1151
513 uValGlyAlaLeuAlaSerAlaLeuSerGlnMetPrLeuArGValay 530
||||||| ||||| ||||| |||||
1152 NGTNGGNCcCYTNGcCmNSNGcCNTNTGTcARcYNTmCCcYTNmGNTNGc 1201
530 alAspLeuTPSerArgGluLeuSerAlaHISGlyAlaLeuAlaTrp 546
||||| ||||| ||||| |||||
1202 TNGATYNTGTGmSGmGmGmGArTYTmSGNCNCARcGNCNGTNGcTGTG 1251
547 PheHISGlnArGArGArgLleLeuGlnGluGlyAlaVallele 563
||||| ||||| ||||| |||||
1252 TYcAYcCNCARmGmGNCARAcYNTNCARGArGNGcGNTNGTNT 1301
563 uLeuPheSerPrOAlaAlaValaGlnCYSGlnGlnTrPLeuGlnLeuG 580
||||| ||||| |||||
1302 NTYNTYmSGNCNGcCNGcCNGcCNTNTGTcYmSGArGTGTcARcAYG 1351
580 LmThrValGluPro.....GlyProHISAlaAlaAlaAlaTrp 593
::: ||||| ||||| ||||| |||||
1352 GNCtMNSNGcCNGcCNGcCNCAYGNCNCARcYNTmGNCmWSN 1401
594 LeuSerCYsValLeuProAspPheLeuGlnGlyArGAlaThrGlyArG 610
::: ||||| ||||| ||||| |||||
1402 YTNmSYNGTNTYNTNCNGAYTYTYTNCARGmNGcCNCNGmNSNTA 1451
610 rValGlyValTTrPheAspGlyLeuLeuHISProAspSerValProSerP 627
||||| ||||| ||||| |||||
1452 YGnTGNGcGTGTYTGAYmGNTYNTNCAYcCNGAYGcGNTNGcGmCY 1501
627 rOpeHArGValaProLeuPheSerLeuProHISGlnLeuProAlaPhe 643
||| ||||| ||||| |||||
1502 TmTmYmGmNCNGTNCNGTmTYAcYTNcCmNSNCARTYTNCNGAYTY 1551
644 LeuAspAlaLeuGlnGlyGlyCYsSerThrSerAlaGlyArGProAlaAs 660
||| ||| |||||
1552 YTNmGNCcCYTNCARcCNCmGNCcCmGmNSNGmNGNTYNTNCARGA 1601
660 pArGValGluArGValThrGlnAlaLeuArGSerAlaLeuAspSerCYs. 676
::: ||||| |||||
1602 RmNGcNCARcArGTYmSGmGNCcCYTNCARcGNCGNTYNGAYmSYNTAY 1651
677ThrSerSerGluAlaProGly 684
::: ||||| |||||
1652 TYcAYcCNCNGcCNGcCNGcCNGcGm 1680

seq_name: gb_hhg:AC024612

seq_documentation_block:

LOCUS AC024612 166620 bp DNA linear HTG 06-MAY-2001

DEFINITION Homo sapiens chromosome 3 clone RP11-481H17 map 3, WORKING DRAFT

SEQUENCE, 22 unordered pieces.

AC024612

AC024612.2 GI:9929740

KEYWORDS HTG; HTGS_PHASEL; HTGS_DRAFT.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 166620)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Bouknight, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choquet, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, R., Dewar, K., Diaz, J.S.,
Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Gratag, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand, P., Grant, G., Hagos, B., Heatford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Laroque, K., Lamazares, R., Landers, T., Lechoczy, J.,
Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,
Meidrum, J., Menus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Olyvar, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (01-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 26, 2000 this sequence version replaced g1:7132995.
All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L1857

Center clone name: 481.H.17

Summary Statistics

Sequencing vector: M13; M7815; 95% of reads

Sequencing vector: Plasmid; n/a; 0-1% of reads

5.051734631771155Chemistry: Dye-terminator Big Dye; 100% of

reads

Assembly program: Phrap; version 0.960731

Consensus quality: 156369 bases at least Q40

Consensus quality: 161786 bases at least Q30

Consensus quality: 163606 bases at least Q20

Insert size: 156000; agarose-fp

Insert size: 164520; sum-of-coverage

Quality coverage: 4.9 in Q20 bases; agarose-fp

Quality co.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 5554: contig of 5554 bp in length
* 5555 5654: gap of 100 bp
* 5655 6844: contig of 1190 bp in length
* 6845 6944: gap of 100 bp
* 6945 8557: contig of 1613 bp in length
* 8558 8657: gap of 100 bp
* 8658 11247: contig of 2590 bp in length
* 11248 11347: gap of 100 bp
* 11348 13427: contig of 2080 bp in length
* 13428 13527: gap of 100 bp

FEATURES

source

13528 17090: contig of 3563 bp in length
* 17091 17190: gap of 100 bp
* 17191 20876: contig of 3686 bp in length
* 20877 20976: gap of 100 bp
* 20977 25280: contig of 4304 bp in length
* 25281 25380: gap of 100 bp
* 25381 29566: contig of 4186 bp in length
* 29567 29666: gap of 100 bp
* 29667 33425: contig of 3759 bp in length
* 33426 33525: gap of 100 bp
* 33526 37279: contig of 3754 bp in length
* 37280 37379: gap of 100 bp
* 37380 44779: contig of 7400 bp in length
* 44780 44879: gap of 100 bp
* 44880 52510: contig of 7631 bp in length
* 52511 52610: gap of 100 bp
* 52611 59797: contig of 7187 bp in length
* 59798 59897: gap of 100 bp
* 59898 66773: contig of 6876 bp in length
* 66774 66873: gap of 100 bp
* 66874 76191: contig of 9318 bp in length
* 76192 76291: gap of 100 bp
* 76292 100510: contig of 24219 bp in length
* 100511 100610: gap of 100 bp
* 100611 111192: contig of 10382 bp in length
* 111193 111292: gap of 100 bp
* 111293 124816: contig of 13524 bp in length
* 124817 124916: gap of 100 bp
* 124917 142057: contig of 17141 bp in length
* 142058 142157: gap of 100 bp
* 142158 157943: contig of 15786 bp in length
* 157944 158044: gap of 100 bp
* 158044 166620: contig of 8577 bp in length.
Location/Qualifiers
1. 166620
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3"
/clone="RP11-481H17"
/clone_lib="RP11 Human Male BAC"
1. 5554
/note="assembly-fragment"
clone_end:SP6
vector_side:left"
5555. 6844
/note="assembly-fragment"
6945. 8557
/note="assembly-fragment"
8658. 11247
/note="assembly-fragment"
11348. 13427
/note="assembly-fragment"
13528. 17090
/note="assembly-fragment"
17191. 20876
/note="assembly-fragment"
20977. 25280
/note="assembly-fragment"
25381. 29566
/note="assembly-fragment"
29667. 33425
/note="assembly-fragment"
33526. 37279
/note="assembly-fragment"
37380. 44779
/note="assembly-fragment"
44880. 52510
/note="assembly-fragment"
52611. 59797
/note="assembly-fragment"
59898. 66773
/note="assembly-fragment"

3360 CGCGCGGGA 73369

Seq_name: gb_pr:AC007783

seq_documentation_block: 168665 bp DNA linear PRI 03-NOV-2001

LOCUS AC007783 Homo sapiens 3 BAC CTFB-140019 (Cal Tech Human BAC Library B)

DEFINITION complete sequence.

ACCESSION AC007783

VERSION AC007783.26 GI:16572920

KEYWORDS HTG.

SOURCE human:

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 168665)

AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alshrocks,S.L., Amaralunga,H.C., Are,J.R., Banks,T., Barberia,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garra,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hayes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hoggins,M., Hollway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Huliyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nicholson,E., Nwokenwo,S., Ogun,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,K., Perez,L., Perez,L., Pichens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,L., Rojals,A., Rojupokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodergren,E., Sonlei,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Swatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczek,R., Woodson,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and Gibbs,R.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 168665)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (11-JUN-1999) Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 168665)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (01-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 4 (bases 1 to 168665)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (03-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 1, 2001 this sequence version replaced gi:16519453.
 INFORMATION: http://www.hgsc.bcm.tmc.edu/or_email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using EPCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT

Summary Statistics

Contig length: 168665
 Phrap values in estimate: 166919
 Average error rate (BCM Phrap estimate): 9.71965e-05
 Fraction of Phrap values less than 40 : 0.0295473
 Number of consensus changing edits: 138
 Number of N's in consensus : 0

Position	Original/Context	Edited/Context
705	ctggagtgca(n)tgccgcgac	ctggagtgca(n)tgccgcgac
732	caatgcagc(n)ccgcctccg	caatgcagc(n)ccgcctccg
737	caagcagc(n)ccgcctccg	caagcagc(n)ccgcctccg
744	cgctcccg(n)ttcagcat	cgctcccg(n)ttcagcat
756	tcatgcat(n)tcgtccca	tcatgcat(n)tcgtccca
759	tgccatntc(n)tgccagc	tgccatntc(n)tgccagc
780	tcacagtag(n)ggactaac	tcacagtag(n)ggactaac
784	agtagcag(n)ctacagtag	agtagcag(n)ctacagtag
802	tgccagcag(n)accccgct	tgccagcag(n)accccgct
811	ccagcccg(n)taattttg	ccagcccg(n)taattttg
841	tagagaggg(n)ttcannnn	tagagaggg(n)ttcannnn
847	cgaggttca(n)annnnnn	cgaggttca(n)annnnnn
849	gggtttca(n)nnnnnn	gggtttca(n)nnnnnn
850	ggtttca(n)nnnnnn	ggtttca(n)nnnnnn
851	tttcannnn(n)nnnnnn	tttcannnn(n)nnnnnn
853	tttcannnn(n)nnnnnn	tttcannnn(n)nnnnnn
855	tttcannnn(n)nnnnnn	tttcannnn(n)nnnnnn
857	tttcannnn(n)nnnnnn	tttcannnn(n)nnnnnn
859	tttcannnn(n)nnnnnn	tttcannnn(n)nnnnnn
860	tttcannnn(n)nnnnnn	tttcannnn(n)nnnnnn
861	tttcannnn(n)nnnnnn	tttcannnn(n)nnnnnn


```

229 PheSerPheLeuLeuTyrLeuArgProValProAspAlaLeuLysSerLe 245
612 ..... 612
245 uTrpTyrLysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrA 262
613 ....CAGAAATTTGTCTGAGAGCCACACTGTAGACTGCTTATGAT 658
262 splenValProCysLeuCysIleGlnValTrpSerLeuGluProAspSer 278
659 TCCTTCCTGCTGATGCTATGAGAGCCCTCACTGACAGAGAGACT 708
279 GluArgValGluPheCysProPheArgGluAspProGlyAlaHis..Ar 294
709 GTAGGCGCAAAAGTGTCTTCAGAGCTGAGCTTGAAGCTTATGCTC 758
294 GAsnLeuTrpHisIleAlaArgLeuArgValLeuSerProGly..... 308
759 AGACTTCGGAGCTCAATACGCTTCACTACTGACAGCAGCAATCA 808
309 ..ValTrpGlnLeuAspAlaProCysCysLeuProGlyLysValThrLeu 324
809 TGTCTATGCTCTGACACTCCGCTGCCACTGAATGAGAGCCCTCCCTC 858
325 CysTrpGlnAlaProAspGlnSerProCysGlnProLeuValProProva 341
859 TGTCTGAGGAGAGAGCCACTCAACCTTCGAAACCTT..... 897
341 LProGlnLysAsnAlaThrValAsnGluProGlnAspPheGlnLeuVala 358
898 .....CCCAAGCCACAGACAGAGAGTGAAGATGATATCTCTGG 940
358 laGly.....HisProAsnLeuCysValGlnValSerThrTrpGlu 371
941 AGAATGTGACTTGCACCCAGCTCTGTTAAGTCTCA...TTTGA 987
372 LysValGlnLeuGlnAlaCys.....SerTr 380
988 AACAGAGCCAGCTTGAAATGTCCCAACAGAGTGGCTCTCCCATCTG 1037
380 palAspSerLeuGlyProPheLysAspMetLeuLeuValGlnMetL 397
1038 GACTGTGAGCATGATAC...CAGGCCAGCAGCTACCGCTTACTTT 1084
397 ysthGlyLeuAsnAsnThrSerValCysAlaLeuGluProSerGlyCys 413
1085 CTTGAGGACATATGCCACTTCACTGCTGAGTGAAGCCAGCTTG 1134
414 .....ThrProLeuProSerMetAlaSerThrArgAlaAlaArgLe 427
1135 GGGCCGGATACCCCATGCTCTCTGTACAGCATCAGCCAGAGCCAGG 1184
427 uGlyGluLeuLeuGlnAsp.....PheArgSerHisG 439
1185 CTCAGTCCACAGTACGCTGACCTCATCCCTTCTGAGGAGGAGA 1234
439 IncysMetGlnLeuTrpAsnAspAspAsnMetGlySerLeuTrpAlaCys 455
1235 ATTGATCTGCTGTGAGAGTCAAT..... 1260
456 ProMetAspLysTyrIleHisArgAlaGlyTrpValLeuVal..... 468
1261 .....GTCCATTTTGGCTGAGAGCAGCTTGTGTCTCTGA 1295
469 .....TrpLeuAlaCysLeuLeuLeuAlaAlaAlaLeuPhePheP 482
1296 TGAAGCCCTTACCTACTGACCTGTCTC..... 1326
482 heLeuLeuLeuLysAspArgAlaGlyAlaAlaArgLysArgThr 498
1327 .....CGGTCCCTAGGCTCCGCTGAGCAAGAGCCA 1356
499 AlaLeuLeuLeuHisSerAlaAspGlyAlaGlyTyrGluArgLeuValG 515

```

```

1357 GTTTACTCTTACATGACGAGCTCAGAGGACAGCAGCGCTGTGG 1406
515 YAlaLeuAlaSerAlaLeuSerGlnMet.....ProLeuArgVala 529
1407 AGCTTTGGCCAACTGCTGCGAGCGCGCTGGAGGTGACCGGAGCTGA 1456
529 lavalAspLeuTrpSerArgArgGluLeuSerAlaHisGlyAlaLeuAla 545
1457 TCGTGATCTCTGGGAGAGGAGGACGATAGCAGCATTTGGACACTGCCG 1506
546 TrpPheHisHisGlnArgArgArgIleLeuGlnGlyGlyValValI 562
1507 TGGCTTTGGGAGCGCGGAGCGCTGGCGCGGAGCAGGAGCAGATGCT 1556
562 eleuLeuPheSerProAlaAlaValAlaGlnCysGlnGlnTrpLeuGln 579
1557 GCTCTGTGAGACTGTGCGGCTCCAGC..... 1584
579 eugInThrValGluProGlyProHisAspAlaLeuAlaAlaTrpLeuSer 595
1585 ..ACCGCTGACGAGGGTGACCGGAGGCTGCCCTTGGCACTTGTG 1632
596 CysValLeuProAspPheLeuGlnGlyArgAlaThrGlyArgTyrValG 612
1633 TGCCTCTCCACGCTCCGCTG.....CTGCT 1658
612 YValTrpPheAspGlyLeuLeuHisProAspSerValProSerProPhe 629
1659 CGCTACTCTCACTGCTGCTGCGCCAAAGTATACATCCCGCGCTCG 1708
629 rGValAlaProLeuPheSerLeuProThrGlnLeuProAlaPheLeuAsp 645
1709 GCGCTGTGCGACGCTACCGGCTGCTGATGACTGCGGCGCTGAGAGA 1758
646 AlaLeuGlnGlyGlyCysSerThrSerAla.....G 656
1759 GCATGTGATGTCAAGCTCCAGCCCTGACCTTCAGCTGAGTCACTTGG 1808
656 YArgProAlaAspArgValGluArgValThrGlnAla.....L 669
1809 GGTACACGGTGTGAAAAACGCTGTGAGCAGTGTCACTGCTGGAAAC 1858
669 euArgSerAlaLeuAspSerCysThrSerSerGlnAlaPro 683
1859 TTGAGCTGCCCAAGATGACTACCAAGGCTCAACCAATGTGCC 1902

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-620-694A-1

seq_documentation_block:
; Sequence 1, Application us/08620694A
; Patent No. 5869286
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 5869286el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; APPLICATION NUMBER: US/08/620,694A
; FILING DATE: 21 MARCH 1996

```


CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/538,765
FILING DATE: 7 AUGUST 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3288 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse
STRAIN: HVS13 receptor
FEATURE:
NAME/KEY: CDS
LOCATION: 121..2715
US-08-620-694A-1

Alignment scores:
Quality: 184.50 Length: 655
Ratio: 0.666 Gaps: 32
Percent Similarity: 42.290 Percent Identity: 20.763

Alignment block:
US-09-899-471-5 x US-08-620-694A-1
Align seg 1/1 to: US-08-620-694A-1 from: 1 to: 3288

198 SerIleGlnSerCysTrp.....ValLeuProTr 207
:::|||||:::|||||
124 GCGATTCGCGCGCTGCTGCGCACGCGTCGTCGCCGCGCGCTGGGATG 173
207 PLeuAsnValSerThrAspGlyAspAsnValLeuLeuThrLeuAspVal 224
|||:::|||||
174 GCTG.....CTTCTGCTGCTGAACGTC 196
224 erGluGlnAspPheSerPheLeuLeuTyrlLeuArgProValPro... 239
197 TGGCCCCGGCGCGCTCCCGCGCTCTCGACTTCGCGCTCCGTC 246
240AspAlaLeu 243
247 TGGCGGAGAGAGGGGCTGAGTGCAGAGTCAGAGATAGTACTTCTCGA 296
243 sSerLeuTrpTyr.....LysAsnLeuThr.....GlyProGlnAsn 256
:::|||||:::|||||
297 TGCAGAGCTGGATCCACCCCAAAACCTGACCCGCTCTCCCAAAAAACA 346
256 LeThrLeuAsn.....HisThrAspLeuValPro 265
|||:::|||||
347 TCTATATCAATTTAGTGTTCCTCTACCCAGCAGAGGAATTAAGTCCCT 396
266 CysLeuCysIleGlnValTrpSerLeuGluProAsp..... 277
|||:::|||||
397 GTGTTCATGATGTTGAG...TGAACCTGCAGACAGATGCCAGATCCCTGA 443
278SerGluArg 280

444 CCTCGAGGGTCAGAGCTGTCCGCTCTGACGCTGAACACCAATGAGCGGC 493
:::|||||
281ValGluPheCysProPheArgGluAspProGlyAlaHisArgAsn 295
|||:::|||||
494 TGTGTGTCAGTTC...CAGTTTGTGTCATGCTCGACAGATCAACGTAAG 540
296 LeuThrPheIleAlaArgLeuArg...ValLeuSerProGlyValTrpG 311
|||:::|||||
541 CGGTGGCGGCTTTCTCTTCAGCCACCTTGTGTAGATCCGTCGC..... 562
311 nLeuAspAlaProCysCysLeuProGlyLysValThrLeuCysTrpGln 328
:::|||||
583CAGAGTATGAGTGAAGTGT..... 603
328 laProAspGlnSerProCysGlnProLeuValProProValProGlnLys 344
:::|||||
604CACCACTGCCGAGAGCCATCCCTGATGG 633
345 AsnAlaThrValAsnGluProGlnAspPheGlnLeuValAlaGlyHisP 361
:::|||||
634 GACCCAAACCAAAATCCA.....AGAT 656
361 roAsnLeuCysValGlnValSerThr..... 369
657 CATCTTGTGCTGCTGACTGTGAGACAGACAGAAATGAAGATGACTACCTAT 706
370TrpGluLysVal 373
707 GCGTGAGCTCAGCAGCCCTTGGATCCACATCATCCTGTGGAGACCTTG 756
373 IGlLeuGlnAlaCysSerTrpAlaAspSerLeuGlyProPheLysAsp 390
:::|||||
757 GACACACAGCATCTGCGAGTGC..... 778
390 spMeLeuLeuValGluMetLysThrGlyLeuAsnAsnThrSerValCys 406
|||:::|||||
779ACTCACCTGT 790
407 AlaLeuGluPro.....SerGlyCysThrProLeu...ProSer 419
:::|||||
791 GGAATGATTCACCCCTTACCAAGCTCTGCTGGAAGTTCTCCGACTCA 840
419 tAlaSerThrArgAlaAlaArgLeuGlyGluLeuLeuGlnAspPhe 436
:::|||||
841 GAGAACCACTGCTGTTGATGTCTTAAACAAATATTTGGCGCCAGCA 890
436 rGSerHisGlnCysMetGlnLeuTrpAsnAspAsnMetGlySerLeu 452
:::|||||
891 AGAAGAATTCATCAGCAGAGGTAATGTCATCACTCACTAACCAAGTTTC 940
453 ..TrpAlaCys..... 455
941 ACTGTGCTGCTGCATCACACGCTGAGTCCAGCCCTCTTCAGACAGTGC 990
455 455
991 CTAAATGACTTTGAGACAGCCTGTGACTGTGCGGCCGAGTAATCTC 1040
456PrometAspLysTyrIleHisArgArgTrpV 466
1041 AAATACACAGATTCACCAAGCAGTTCAGACACTCAAT...CCCGTGG 1087
466 al.....LeuValTrpLeuAlaCysLeuLeuAlaAlaLeuPhe 480
|||:::|||||
1088 TGTATGGCTTCATCATCACTCATCGCATTCGCTGCTGGATCTGATC 1137
481 PhePheLeu.....LeuLeuLysLysAspArgArg 491
|||:::|||||
1138 GTGCTGATCATCTGTATGACCTGAGGCTTTCGGCGGATCAAGAGAA 1187
491 sAlaAlaArgGlySer..... 496

CITY: Seattle

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

243 sSerLeuTrpTyr.....LysAsnLeuThr.....GlyProGlnAsn 256
    ::  |||  ||||| |||||  ::|||::|||

```


297 TGACGCTGGATCCACCCCAAAACCTGACCCCGCTTCCCAAAAAACA 346
 256 IeThrLeuAsn.....HisThraspleValPro 265
 347 TCTATATCAATCTTAGTGTTCCTCCTACCCAGCAGGAGATATGTCTCCT 396
 266 CysLeuCysIleGlnValTrpSerLeuGluProAsp..... 277
 397 GTGTTCATGTGAG...TGACCCCTGACAGACAGATCCAGCATCTGTGA 443
 278SerGluArg. 280
 444 CCTGAGGGTGCAGAGCTGTCCGTCTGACGTGAACACCAATGACGGGC 493
 281ValGluPheCysProPheArgGluAspProGlyAlaHisArgAsn 295
 494 TGTGTCAAGTTC...CAGTTTCTCTCATGCTGCAGCATCACCGTAAG 540
 296 LeuTrpHisIleAlaArgLeuArg...ValLeuSerProGlyValTrpG1 311
 541 CGGTGGCGGTTTCTTCCAGCACTTGTGTGATGATCCCTGGC..... 582
 311 nLeuAspAlaProCysCysLeuProGlyValTrpLeuCysTrpGlnA 328
 583CAGAGATATGAGTACTGTT..... 603
 328 LaProAspGlnSerProCysGlnProLeuValProValProGlnLys 344
 604CACCACTGCCGAGAGCCATCCCTGATGGC 633
 345 Asn..AlaThrValAsnGluProGlnAspPheGlnLeuValAlaGlyHisP 361
 634 GACCCAAACCAAAATCCA.....AGAT 656
 361 roAsnLeuCysValGlnValSerThr..... 369
 657 CATCTTGTGCTGACTGTGAGACAGCAAGATGAAGTACTACTCAT 706
 370TrpGluLysVa 373
 707 GCGTGAGCTCAGGAGCCTTTCGATCCCAACATCACTGTGGAGACTTG 756
 373 IGIleuGlnAlaCysSerTrpAlaAspSerLeuGlyProPheLysAspA 390
 757 GACACACAGCATCTCGGAGTGG..... 778
 390 spMetLeuLeuValGluMetLysThrGlyLeuAsnAsnThrSerValCys 406
 779ACTTCACCCCTGT 790
 407 AlaLeuGluPro.....SerGlyCysThrProLeu...ProSerme 419
 791 GGAATGAAATCCACCCCTACACGAGTCTCTGGAAGTTTCTCCGACTCA 840
 419 talSerThrArgAlaAlaArgLeuGlyGluLeuLeuGlnAspPhea 436
 841 GAGAACCAACACTCTTGTGATCTCTTAACAAATATTTCCGCCAGCA 890
 436 rGSerHisGlnCysMetGlnLeuTrpAsnAspAspAsnMetGlySerLeu 452
 891 AGAAGAATTCATCAGCGAGCTAATGTACATTCACCTAAGCAAGTTTC 940
 453TrpAlaCys..... 455
 941 ACTGTGTCTGCCATCACCACGTGACAGTCCAGCCCTTCTTCACAGACTGC 990
 455 455
 991 CTAAATGACTGTTTGAGACAGCTGTGACTGTGCCCTGCCAGTAATCTC 1040
 456PrometAspLysTyrIleHisArgArgTyr 466
 1041 AAATACACAGTTCACCAAGCCAGTTGACAGACTACATTT...CCCCTGTGG 1087

466 al.....LeuValTrpLeuAlaCysLeuLeuLeuAlaAlaLeuPhe 480
 1088 TGTATGCTTCATCATCACTCATTCGCCATTCTCTGTGTGGATCTGTCAATC 1137
 481 PhePheLeu.....LeuLeuLysLysAspArgArg 491
 1138 GTGTGATTCATCTGTATGACCTCGAGAGCTTCTGTGCCGCCGATCAAGAA 1187
 491 saAlaAlaArgGlySer..... 496
 1188 ACATGTGATGATCTCCAAATCAATGATGCATTTGCCCTGACAGACTGA 1237
 497ArgTrpAlaLeuLeuHisSerAlaAsp 506
 1238 CTCGCCCAACCCCTGAGGCCAGGAAGTGTGATGCTTACTGTGGCCGAC 1287
 507 GLYAlaGlyTyrGluArgLeuValGlyAlaLeuAlaSerAlaLeu...Se 522
 1288 CACCCCTCATGTGAGAGGTGTCTCTAAAGTTCCGCCAGTCTCTGTATCAC 1337
 522 rGlnMetProLeuArgValAlaValAspLeuTrpSerArgArgGluLeu 539
 1338 TGCCCTGTGGCAGCTGAGTAGCCCTTGACCTCTCGAGAGAGCAGTTATCT 1387
 539 eAlaHisGlyAlaLeuAlaTrpPheHisGlnArgArgArgIleLeu 555
 1388 CTGAGTGTGGGGTATGATCACTGTGGAGCCGACAGAAAGCAGATGTGTG 1437
 556 GlnGluGlyValValIleLeuLeuPheSer..... 566
 1438 GAGAGCAATCCAAATCATCATCTCTGTGTCCGAGGACCAACCAAGCAMA 1487
 567ProAlaAlaValAlaGlnCysG 574
 1488 GTGGAAGCATCTTGTGGTGGCTGAGCCCTGTCTCCAGTCACTGTGTG 1537
 574 IncInTrpLeuGlnLeuGlnThrValGluProGlyProHisAspAlaLeu 590
 1538 ACCACTGG.....AGCCTGCT...GGGACCTTTTC 1566
 591 AlaAlaTrpLeuSerCysValLeuProAspPheLeuGlnGlyArgAlaTh 607
 1567 ACTGACGACATGACATGATCTCTGCCAGACTTCAAGAGCCAGCCCTGCTT 1616
 607 rGlyArgTyrValGlyValTyrPheAspGlyLeuLeuHisProAspSerV 624
 1617 CGGCACCTACGTGTGTCTCTACAGTCCAGTCTGTAGTACAGAGGATG 1666
 624 alProSerProPheArgValAlaProLeuPheSerLeuPro...ThrGlnLe 640
 1667 TCCCGACCTCTTCAATCATCACTCCAGGTACCCACTCATGACAGATTT 1716
 640 uproAlaPheLeuAspAlaLeuGln.....GlyG 650
 1717 GAGAGAGTTACTTCCGATCCAGACCTGAGATGTTTGAACCGCGCG 1766
 650 LysSer.....ThSerAlaGlyArgProAspArgValGlu 663
 1767 GATGACACCATGTGACAGAGCTCAGAGGACAAATTAATCTGCGAGCCCTA 1816
 664 ArgValThrGlnAlaLeuArgSerAlaLeuAspSerCysThrSerSerse 680
 1817 GTGGCGGCGACGTCAAGAGGCTGTGTAGTTCCAGAGATGGCAACC 1866
 680 rGluAlaPro 683
 1867 CAGTGCCTCG 1876

seq_name: /cgn2.6/ptdata/2/lna/5A.COMB.seq:US-09-022-696-1
 seq_documentation_block:
 ; Sequence 1, Application us/09022696

224	erglgluglnlaspheSerPheLeuLeuTyrLeuArgProValPro...	239
197	TGGCCCGGGCCGGGCTCCCGCGCGCTCTCGACTTCGGCGCGGTC	246
240AspIleLeu	243
247	TGCCGCGAGAGGGGTGAGCTGCAGAGTCAGAGATAGTACTTGTGTGA	256
243	sSerLeuPyr.....LysAsnLeuThr.....GlyProGlnAsn	256
297	TGACAGCTGGANTCCACCCCAAAACTGACCCCGCTTCCCAAAAACA	346
256	LeuThrLeuAsn.....HisThrAspLeuValPro	265
347	TCTATATCAATCTAGTGTTCCTTACCCACACAGCAATATGATGTCCT	366
266	CysLeuCysIleGlnValThrSerLeuGlnProAsp.....	277
397	GTGTTCAGTGTGAG...TGGACCTGCAGACAGATGCCAGCATCTGTA	443
278SerGluArg	280
444	CTTGAGGGTGCAGAGCTGTCCGCTCTGCAGCTGAACACCAAGACGGC	493
281	ValGluPheCysProPheArgGluAspProGluValHisArgAsn	295
494	TGTGTGCAAGTTC..CAGTTCTGTCTCATGCTGCAGAGTACCGTAA	540
296	LeuThrHisIleLeuArgLeuArg...ValLeuSerProGluValThrG	311
541	CGGTGGCGGTTCCTTGACGACCATTTGTGTGATGATCTGGC.....	582
311	LeuAspAlaProCysCysLeuProGluValThrValThrLeuCysThrGln	328
583CAGAGTATGAAATGATGTT.....	603
328	IaProAspGlnSerProCysGlnProLeuValProProValProGlnLys	344
604CACACCTGCCAGAGCCATCCTGATGGC	633
345	AsnAlaThrValAsnGluProGlnAspPheGlnLeuValAlaGluHisP	361
634	GACCAACCAACCAATCA...AGAT	656
361	roAsnLeuValGlnValSerThr.....	369
657	CATTTGTGCTGCTACTGTGAGGACAGACAGATGAATGACTACTCAT	706
370TyrGluLysVal	373
707	GCGTAGAGTCAAGGACGCTTTGGATCCCAACATCACTGTGGAGACTTG	756
773	IgluLeuGlnAlaCysSerThrPalaAspSerLeuGluYrProPheLysAsp	390
757	GACACACAGCATCTGCGAGTGG.....	778
390	spMetLeuLeuValGluMetLysThrGluLeuAsnAspThrSerValCys	406
779ACTTCACCTGT	790
407	AlaLeuGluPro.....SerGlyCysThrProLeu...ProSer	419
791	GGATGTGAATCCACCCCTCACAGAGTCCGTGGAAAGTTTCCCACTCA	840
419	AlaSerThrArgAlaAlaArgLeuGluLeuLeuLeuGlnAspPheA	436
841	GAGAACCCACAGCTGCTTGATGTCGTTAAACAAATTTGGCCCGCCAGCA	890
436	rgSerHisGlnCysMetGlnLeuThrPAsnAspAspAsnMetLysSerLeu	452
891	AGAAGAAATTCATGACGAGGCTAATGTCACTTCACTTCAAGCAAGTTTC	940


```

453 ..TTPalacys..... 455
941 ACTGTGCTGCTGCATCCACGACGTGACGAGCCCTTCTTACAGACGCTGC 990
455 ..... 455
991 CTAAATGACTTTGAGACAGCGCTGACTGTGCCCTGCCCACTATATCTC 1040
456 .....PrometaspIystrilHHisargargtrpv 466
1041 AATATACACAGCTTCCCAAGCCAGTGGACAGACTACAT...CCCTGTGGG 1087
466 al.....LeuValtripleuAlaCysLeuLeuLeuAlaAlaaleuPhe 480
1088 TGTATGCGCTTACACACTCATCGCCATCTCTGCTGGGATCTGTCATC 1137
481 PhePheLeu.....LeuLeuLysLysAsparargly 491
1138 GTGCTGATCATCTGATATGACCTGGAGGCTTCTGGCCCGCATCAAGAA 1187
491 salAlaIaarglyser..... 496
1188 ACATGTGATGATCTCCAAATGATGGATCTTCCCGTAGACAGACTGA 1237
497 .....ArgThrAlaLeuLeuHisSerAlaasp 506
1238 CTCGCCACCCCTGAGGCCCAAGAGGTCGTGATCTGCTACTCGCCGAC 1287
507 G1yAlaIaIyrgIuargLeuValIyAlaLeuAlaSerAlaLeu...Se 522
1288 CACCCCTCTATGTGAGGCTGCTTAAGTGGCCAGTCTGATCAC 1337
522 rgImetProLeuArgValAlaAlaAspLeuTrpSerArgArgIuleus 539
1338 TGCCCTGTGCGACGAGTACGCTTGTGACCTCTGGAAGACAGATATCT 1387
539 erAlaHisGlyAlaLeuAlaTrpPheHisHisGlnArgArgIleLeu 555
1388 CTGAGGTGGGGTCACTGACCTGGGAGCGGACAGAAAGCAGAGATGCTG 1437
556 GlnGluGlyValIyAlaIleLeuLeuPheSer..... 566
1438 GAGAGCACTCCAAATCATCTGTGTTCCGAGGACCCAGCAAA 1487
567 .....ProAlaAlaValAlaGlnCysG 574
1488 GTGGAAAGCTATCTTGGGTGGCGTGAGCCTCTGTCCAGCTACGGGTG 1537
574 InGlnTrpLeuGlnLeuGlnThrValGluProGlyProHisAspAlaLeu 590
1538 ACCACTGCT.....AAGCTGCT...GGGACCTTTTC 1566
591 AlaAlaIleTrpLeuSerCysValLeuProAspPheLeuGlnIyAlaAlaTrp 607
1567 ACTGCACGCCATGACATGATCTCTGCCAGATCAAGAGCCACGCTT 1616
607 rGlyArgTyValGlyValIyTrpPheAspGlyLeuLeuHisProAspSer 624
1617 CGGCACTAGCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1666
624 alProSerProPheArgValAlaIleProLeuPheSerLeuPro...ThrGln 640
1667 TCCCGCAGCTTCTTCAACATCACTCCAGTACCCACTATGACAGATTT 1716
640 uProAlaPheLeuAspAlaLeuGln.....GlyG 650
1717 GAGGAGGTTTACTTCCGATCCAGACCTGAGAGTGTGAAACCCGCGC 1766
650 LysSer.....ThrSerAlaGlyArgProAlaAspArgValGlu 663
1767 GATGCACCAATCTGACAGAGCTCACAGGGACAAATTACCTGCAGAGCCCTA 1816
664 ArgValThrGlnAlaLeuArgSerAlaLeuAspSerCysThrSerSer 680

```

```

1817 GTGGCCGACGCTCAAGAGAGCTGTGCTTGTAGTTCAGAGAGTGCAGAAACC 1866
680 rGluAlaPro 683
1867 CAGTGCCTCG 1876

```

seq_name: /cgn2_6/plodata/2/lna/6A_COMB.seq:us-08-978-773-1

seq_documentation_block:

Sequence 1, Application US/08978773

Patent No. 6083906

GENERAL INFORMATION:

APPLICANT: Trout, Anthony

TITLE OF INVENTION: Method of Regulating Nitric Oxide Production

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESS: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple PowerMacintosh

OPERATING SYSTEM: Apple Operating System 7.5.5

SOFTWARE: Microsoft Word for PowerMacintosh, Version 6.0.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/978,773

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 60/052,525

FILING DATE: 27 NOVEMBER 1996

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,693

REFERENCE/DOCKET NUMBER: 2623-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 587-0430

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3288 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Mouse

IMMEDIATE SOURCE:

CLONE: IL-17 receptor

FEATURE:

NAME/KEY: CDS

LOCATION: 121..2712

US-08-978-773-1

alignment_scores:

Quality: 184.50

Ratio: 0.666

Percent Similarity: 42.290

Percent Identity: 20.763

alignment_block:

US-09-899-471-5 x US-08-978-773-1 ..

Align seg 1/1

to: US-08-978-773-1

from: 1

to: 3288

198 SerIleGlnSerCysTrp.....ValLeuProTr 207


```

124 GCGATTCGGCGCTGCTGGCCACGGGTCTCCCGCGCGCGCTGGGATG 173
:::|||||
207 pLeuAsnValSerThrAspGlyAspAsnValLeuLeuThrLeuAspValS 224
|||||
174 GCTG.....CTTCTGCTGTGAACGTTG 196
224 eArgLIGLInAspPheSerPheLeuLeuTyLeuArgProValPro... 239
|||
197 TGGCGCCGGCGCGCGCTCCCGCGCTCTCGACTTCCCGCTCGGTC 246
240 .....AspAlaLeuLy 243
247 TGGCGCAGAGAGGGCTGAGTCGAGATCAAGATAGTACTGTCTGGA 296
243 sSerLeuThrPyr.....LysAsnLeuThr.....GlyProGlnAsnI 256
:::|||||
297 TGACAGCTGATCCACCCCAAAACCTGACCCCGTCTCCCAAAACA 346
256 LeuThrLeuAsn.....HisThrAspLeuValPro 265
|||
347 TCTATATCAATCTTAGTGTTCCTCTACCCAGCAGCGAGAAATTAGTCC 396
266 CysLeuCysIleGlnValTrpSerLeuGluProAsp..... 277
|||
397 GTGTTGCATGTTGAG...TGACCCCTGCAGACAGATGCCAGATCCTGTA 443
278 .....SerGluArg. 280
444 CCTGAGGGGTGCAGAGCTGCTCCGCTGCAGCTGAACACCAATGACGGC 493
281 .....ValGluPheCysProPheArgGluAspProGluAlaHisArgAsn 295
|||||
494 TGTGTGTCAAGTTC...CAGTTTGTCTCATGCTGCAGCATCACCCTAG 540
296 LeuThrPheIleAlaArgLeuArg...ValLeuSerProGluValTrpGlu 311
|||||
541 CGGTGGGGGTTTCTCTCAGGCACATTGGTAGATCCGTGGC..... 582
311 mLeuAspAlaProCysCysLeuProGluValValThrLeuCysTrpGlu 338
583 .....CAGGAGTATGCAAGTACTGTT..... 603
328 laProAspGlnSerProCysGlnProLeuValProProValProGlnLyS 344
604 .....CACCACTGCGCAGAGCCCATCCCTGATGGG 633
345 Asn..AlaThrValAsnGluProGlnAspPheGlnLeuValAlaGluHisP 361
:::|||||
634 GACCCCAACCCCAAAATCCA.....AGAT 656
361 roAsnLeuCysValGlnValSerThr..... 369
657 CATCTTGTGCTGACTGTGTAGGACAGACAGATGAGATGACTACTCTCAT 706
370 .....TrpGluLysVa 373
707 GCGTAGCTCAGGACGCTTGGATCCCAACATCAGTGTGGAGACTTGG 756
373 IGlnLeuGlnAlaCysSerTrpAlaAspSerLeuGluPropheLysAspA 390
:::|||||
757 GACACACAGCATCTGCGAGTGG..... 778
390 spMetLeuLeuValGluMetLysThrGlyLeuAsnAsnThrSerValCys 406
|||||
779 .....ACTTCACCGCTGT 790
407 AlaLeuGluPro.....SerGlyCysThrProLeu...ProSerme 419
|||||
791 GGAATGATTCACACCCCTACAGGTCTCGTGGAAAGTTTCTCCGACTCA 840
419 tAlaSerThrArgAlaAlaArgLeuGluGlnLeuLeuGlnAspPheA 436
:::|||||
841 GAGAACCCACAGCTGCTTGTGATGTCTTAACAATATTTTGGCCCAAGCA 890
436 rGSerHisGlnCysMetGlnLeuTrpAsnAspAspAsnMetGlySerLeu 452
|||||
891 AGAAGATTCATCAGGAGGAGCTAATATGCATTCACCTCTAAGCAAGTTTC 940
453 ..TrpAlaCys..... 455
941 ACTGGGTGGCCATCAACACAGTCGAGGTCCAGCCCTTTCTCAGCAGTGC 990
455 ..... 455
991 CTAAATGACTGTTTGAGACACGCTGTGACTGTGCCCTCCAGATATTC 1040
456 .....PrometAspLysTrpIleHisArgArgTrp 466
1041 AAATACACAGTTCCCAAGCCAGTTGCAGACTACATT...CCCTGTGG 1087
466 al.....LeuValTrpPheAlaCysLeuLeuLeuAlaAlaLeuPhe 480
|||
1088 TGTATGGCTCATCAACACACTCATGCGCATTCGTGCTGGGGATCTGTCA 1137
481 PhePheLeu.....LeuLeuLysLysAspArgArg 491
1138 GTGCTATATCATCTGTATGACCTGAGGCTTCTGCGCCGATCAAGAA 1187
491 saAlaAlaArgGlySer..... 496
1188 ACATGCTGATGACTCCAAATCAATGCGATCTTCCCGCTACACACCTGA 1237
497 .....ArgThrAlaLeuLeuLeuHisSerAlaAsp 506
1238 CTCCCCACCCCTCAGAGCCCAAGAGGTGTGATCGTCTACTCGGCGCAG 1287
507 GLyAlaGlyTrpGluArgLeuValGlyAlaLeuAlaSerAlaLeu...Se 522
1288 CACCCCTCTATGTGAGGGGTGTCTAAGTTGCCCCAGTTCTGATCAG 1337
522 rGlnMetProLeuArgValAlaValAspLeuTrpSerArgArgGluLeuS 539
1338 TCCCTGTGACACTGAAGTAGCCCTTGACCTCTGAGAGAGAGGTTTACT 1387
539 eAlaHisGlyAlaLeuAlaTrpPheHisGlnAlaGluArgGlyLeuLeu 555
|||
1388 CTGAGGTGGGGCTCATGACCTGGGTGAGCCGACAGAGGAGATGGTG 1437
556 GlnGluGlyGlyValValIleLeuLeuPheSer..... 566
1438 GAGAGCACTCCAAATCAATCATCTGTGTGCCGAGGCCAACGAAA 1487
567 .....ProAlaAlaValAlaGlnCysG 574
1488 GTGAAAGCTATCTTGGGTGGCTGAGCTGCTGTCAGACTAGCGGTGTG 1537
574 InGlnTrpLeuGlnLeuGlnThrValGluProGluPheHisAspAlaLeu 590
1538 ACCACTGG.....AAGCTGCT...GGGACCTTTTC 1566
591 AlaAlaTrpLeuSerCysValLeuProAspPheLeuGlnGluArgAla 607
1567 ACTGCAAGCATGAACATGATCCCTGCCAGCTTCAAGAGGCGAGCTGCTT 1616
607 rGlyAlaGlyValGlyValTrpPheAspGlyLeuLeuHisProAspSer 624
1617 CGGCACCTGACTGTTGTTGCTACTTCAAGTGCATCTGATGAGAGGATG 1666
624 alProSerProPheArgValAlaProLeuPheSerLeuPro..ThrGlnLe 640
1667 TCCCGACCTCTTAACATCACTCCAGTACCCACTCATGAGACAGATT 1716
640 uProAlaPheLeuAspAlaLeuGln.....GlyG 650
1717 GAGGAGGTTTACTTCCGAGATCCAGAGACTGGAGATGTTTAACCCGCGC 1766

```


LOCATION: 121..2715

373 IGIInEugInaIaCysSerTTPaIaasPserLeuGIyProPhelysaspA 390

757 GACACACAGCATCTGCAGATG..... 778
390 spkMetLeuValGluMetIysThrIleuAsnAspThrSerValCys 406
779ACTTACACCTCT 790
407 AlaLeuGluPro.....SerGlyCysThrProLeu...ProSerIle 419
791 GGATATATTCACACCCCTTACAGAGCTCTGCAAGAGTTTCTCCACATCA 840
419 talaserThrArgAlaIleAlaIleArgLeuGlyGluIleuLeuGlnAspPhe 436
841 GAGACACACAGCTCTTGTGATGCTTAACAAATATTGGCCCAAGCA 890
436 rSerHisGlnCysMetGlnLeuThrPasnAspAspMetGlySerLeu 452
891 AGAAGATTCATCAGCAGAGCTAATGTCACATTCCTTAAGCAAGTTTC 940
453 ..TrpAlaCys..... 455
941 ACTGCTGCTCCATCACACAGTGCAGTCCAGCCCTTCTTACAGAGCTGC 990
455 455
991 CTAAATGACTGTTGAGACACGCTGTACTGCTGCCCTGCCAGTATCTTC 1040
456PrometAspIysTyrIleHisArgArgTyrPv 466
1041 AATATCACAGATTCACCAAGCATTCGACAGTACAT...CCCTGTGGG 1087
466 al.....LeuValTrpLeuAlaCysLeuLeuAlaIleAlaLeuPhe 480
1088 TGTATGCTTCATCAGACATCATCCCATCTGCTGGGATCTGTGCATC 1137
481 PhePheLeu.....LeuLeuIysLysAspArgArg 491
1138 GTCCTATCATCTCTATGACTGCTGAGGCTTCTGCGCCGATCAAGAGA 1187
491 salalaAlaArgLys..... 496
1188 ACATGGATGATGACTCCAAATCATGTCATCTTGCCTGACAGACTGA 1237
497ArgThrAlaLeuLeuHisSerAlaAsp 506
1238 CTTCCCTCCACCTGAGGCCAGGAAGTCTGATGCTGATCTGCGCGAC 1287
507 GtAlaGlyTyrGluArgLeuValGlyAlaLeuAlaSerAlaLeu...se 522
1288 CACCCCTCTATGAGAGGTGCTCTTAAGTTCCGCCAGTCTCTGTATC 1337
522 rGlnMetProLeuArgValAlaValaAspLeuTrpSerArgArgGluLeu 539
1338 TGCTCTGAGCATGAAGTACCTTACCTCTGGAAGAGAGGTATATCT 1387
539 exAlaHisGlyAlaLeuAlaTrpPheHisGlnArgArgGlyLeu 555
1388 CTGAGTGGGGGTGATGACCTGGGTGAGCCACAGAGAGAGATGTG 1437
556 GlnGluGlyValValaIleLeuLeuPheSer..... 566
1438 GAGAGCACTCCAAATCATCTCTGTCTTCCAGAGCAGCCCAAGCAA 1487
567ProAlaAlaValaIleGlnCysG 574
1488 GTGGAAGTATCTTGGGTGGCTGAGCTCTGTCAGTACAGGTATG 1537
574 InGlnTrpLeuGlnLeuGlnThrValGluProGlyProHisAspAlaLeu 590
1538 ACCACTGG.....AAGCTGTCT...GGGACCTTTTC 1566
591 AlAlaTrpLeuSerCysValLeuProAspPheLeuGlnGlyArgAlaTh 607

1567 ACTGACGACATGACATATCTCTGCAGACTTCAGAGGACGACCTGCTT 1616
607 rGlyArgTyrValGlyValTyrPheAspGlyLeuLeuHisProAspSer 624
1617 CGGACACATCTGCTTCTCTACTGAGTCATCTGATGAGAGGATG 1666
624 alProSerProPheArgValAlaProLeuPheSerLeuPro ThrGlnIle 640
1667 TCCCGACCTCTTCAACATCACCCTCAGGTACCCACTCATGACAGATTT 1716
640 uProAlaPheLeuAspAlaLeuGln.....GlyG 650
1717 GAGAGGTTTACTTCCGATCCAGGACCTGAGATTTTGAACCGGCG 1766
650 lYcYser.....ThrSerAlaGlyArgProAlaAspArgValGlu 663
1767 GATGACACATGTCAGAGAGCTCAGAGGAGCAATTCCTGACAGACTTA 1816
664 ArgValThrGlnAlaLeuArgSerAlaLeuAspSerCysThrSerSer 680
1817 GTGGCGGACGCTCAAGAGGCTGTGCTTAGTTTCCAGGAGTGCACAAAC 1866
680 rGluAlaPro 683
1867 CAGTCCCGCG 1876
seq_name: /cgn2.6/ptodata/2/ina/6A_COMB.seq:US-09-022-260-1
seq_documentation_block:
; Sequence 1, Application us/09022260
; Patent No. 6100235
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: NO. 6100235el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,260
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/620,694
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USN 08/410,535
; FILING DATE: 23 MARCH 1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2617-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Mouse
 STRAIN: HVS13 receptor
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 121..2715
 US-09-022-260-1

Alignment scores:
 Quality: 184.50 Length: 655
 Ratio: 0.666 Gaps: 32
 Percent Similarity: 42.290 Percent Identity: 20.763

Alignment block:
 US-09-899-471-5 x US-09-022-260-1 ..

Align seg 1/1 to: US-09-022-260-1 from: 1 to: 3288

198 SerIleGlnSerCysTrp.....ValLeuProTr 207
 124 GCCATTGCGGCTGTGGCCAGCGGTCTCCCGGCGCGCTGGGATG 173
 207 PleuAsnValSerThrAspGlyAspAsnValLeuLeuThrLeuAspValS 224
 174 GCTG.....CTTCTGCTGCTGAAGCTTC 196
 224 erGlugInGlnAspPheSerPheLeuLeuTyrLeuArgProValPro... 239
 197 TGGCCCCGGCGCGCTCCCGCGCTCTCGACTCCCGGCTCCGGTC 246
 240AspAlaLeu 243
 247 TGGCCGAGGAGGGGTGAGCTGACAGTCAAGAAATAGTACTGTCTGA 296
 243 sSerLeuTrpTyr.....LysAsnLeuThr.....GlyProGlnAsnI 256
 297 TGACAGCTGATGCACCCCAAAACCTGACCCCGCTTCCCAAAAACA 346
 256 IerHLeuAsn.....HisThrAspLeuValPro 265
 347 TCTATATCAATCTTAGTGTTCCTCTACCCAGACGAGAAATTAGTCCCT 396
 266 CysLeuCysIleGlnValTrpSerLeuGluProAsp..... 277
 397 GTCTTGATGTGTGAG...TGGACCTCGACAGATGCCAGCATCCTGTA 443
 278SerGluArg 280
 444 CCGGAGGAGTGCAGAGCTCCGCTGCTGACAGTGAACACATGACGGC 493
 281ValGluPheCysProPheArgGluAspProGlyAlaHisArgAsn 295
 494 TGTGTGCAGATTC...CAGTTCTGTCCATGCTGACAGATCAGCCTGAG 540
 296 LeuTrpHisIleAlaArgLeuArg...ValLeuSerProGlyValTrpG 311
 541 CGGTGGGGGTTTCTTCAGCCACTTGTGTGATGATCCGTGGC..... 582
 311 nLeuAspAlaProCysLeuProGlyLysValThrLeuCysTrpGlnA 328
 583CAGGATATGAAGTACTGTT..... 603
 328 IaProAspGlnSerProCysGlnProLeuValProProValProGlnLys 344
 604CACCACCTGCCGAAGCCATCCCTGATGG 633
 345 AsnAlaThrValAsnGluProGlnAspPheGlnLeuValAlaGlnHisP 361

634 GACCCAAACCAAAATCCA.....AGAT 656
 361 roAsnLeuCysValGlnValSerThr..... 369
 657 CATCTTTGTGCTGACTGTGAGACACAGCAAGATGATGACTACTCCAT 706
 370TrpGluLysVal 373
 707 GCGTGTGCTGACGACGCTTTGGATCCCAATCATCTGAGAGACCTTG 756
 373 IeInLeuGlnAlaCysSerThrPalaAspSerLeuGlyProPheLysAsp 390
 757 GACACACAGCATCTCGAGTGG..... 778
 390 spMetLeuLeuValGluMetLysThrGlyLeuAsnAsnThrSerValCys 406
 779ACTTCACTCTGT 790
 407 AlaLeuGluPro.....SerGlyCysThrProLeu...ProSerMe 419
 791 GGAATGATTCACACCCCTTACACAGTCTCTGGAAGTTCTCCGACTCA 840
 419 ValSerThrArgAlaAlaArgLeuGluGluLeuLeuGlnAspPheA 436
 841 GAGAACACACAGCTGCTTGTGATGCTGTTAAACAAATATTTGGCCAGGCA 890
 436 rgSerHisGlnCysMetGlnLeuTrpAsnAspAspMetGlySerLeu 452
 891 AGAAGATTCATCAGCAGAGCTAATGTCATCATTCATCTAAGCAAGTTTC 940
 453 ..TrpAlaCys..... 455
 941 ACTGTGCTGATCATCACCACAGCTGCAGGCTCCCTTCAGACAGCTGC 990
 455 455
 991 CTAAATGACTGTTTGTGAGACAGCTGTGACTGTGCCCTGCCAGTAATCTC 1040
 456PromelasplustyrIleHisArgLysTrp 466
 1041 AAATACACAGTCTCCACAGCCAGTTCACAGCTACATTT...CCCTGTGG 1087
 466 al.....LeuValTrpLeuAlaCysLeuLeuLeuAlaAlaLeuPhe 480
 1088 TGTATGGCTCATCATCAGCTCATGCGCATCTGCTGTGGGATCTGTATC 1137
 481 PhePheLeu.....LeuLeuLysLysAspArgArgly 491
 1138 GTCTGTATCATCTGTATGACCTGAGAGCTTTCTGCGCGCATCAAGAA 1187
 491 sAlaIlaArgLysSer..... 496
 1188 ACATGTGATGACTCCAAATCAATGACATCTGCCCTGACAGACCTGA 1237
 497ArgThrAlaLeuLeuLeuHisSerAlaAsp 506
 1238 CTCCGCCACCCCTGAGGCCAGAGAGTCTGATGCTGTCTGTGGCGGAC 1287
 507 GlyAlaGlyTyrGluArgLeuValGlyAlaLeuAlaSerAlaLeu...Se 522
 1288 CACCCCTCTATGTGTGAGGTGTCTAAAGTTGGCCAGTCTCTGATCAC 1337
 522 rGlnMetProLeuArgValAlaValAspLeuTrpSerArgArgGluLeuS 539
 1338 TGCCTGTGACAGTGAAGTACCTTGAACCTCTGAGACAGAGTTATCT 1387
 539 eAlaHisGlyAlaLeuAlaTrpPheHisGlnArgArgArgIleLeu 555
 1388 CTGAGGTGGGGTCACTGACTGTGGTGAAGCCGACAGAAAGAGAGATGTG 1437
 556 GlnGluGlyGlyValValIleLeuLeuPheSer..... 566
 1438 GAGAGCAACTCCAAATCATCATCTGTGTCCGAGGACACCCAAAGCAA 1487

ATTORNEY/AGENT INFORMATION:

[illegible]


```

311 nleuaspalaProCysCysLeuProGlyValThrLeuCysThrGlna 328
583 .....CAGAGATGAGTACGTGT ..... 603
328 laProaspInserProCysGlnProLeuValProProValProGlnLys 344
604 .....CACCACCTGCCGAAGCCCATCCCTGATGGG 633
345 AsnAlaThrValAsnGluProGlnAspPheGlnLeuValAlaGlyHisP 361
634 GACCCAAACCACCAATCCA.....AGAT 656
361 roAsnLeuCysValGlnValSerThr..... 369
657 CATCTTGTGCTGACTGTGAGACAGCAAGATGATGATGATGATCAT 706
370 .....TrrGlyLysVa 373
707 GCGTAGCTCAGGACGCTTTGGGATCCCAATCATCTGTGGAGACCTTG 756
373 GlnLeuGlnAlaCysSerThrLysAspSerLeuGlyProPheLysAsp 390
757 GACACACAGCATCTGGAGTGG..... 778
390 spMetLeuLeuValGluMetLysThrGlyLeuAsnAsnThrSerValCys 406
779 .....ACTTCACCTGT 790
407 AlaLeuGluPro.....SerGlyCysThrProLeu..ProSerMe 419
791 GGAATGAATCCACCCCTTACACAGCTCTGTGGAAGTTCTCCGACTCA 840
419 TalSerThrArgAlaAlaArgLeuGluGluLeuLeuGlnAspPhe 436
841 GAGAACACACAGCTGTGTGATGCTGTTAAACAATTTTGGCCCGACGA 890
436 rgSerHisGlnCysMetGlnLeuThrAsnAspAspMetGlySerLeu 452
891 AGAAGAAATTCATCAGCGAGCTAATGCTACATTCCTAAGCAAGTTTC 940
453 ..TrpAlaCys..... 455
941 ACTGTGTGCTGATCCACACAGCTGACGTCACGCCCTTCTTACAGAGCTGC 990
455 ..... 455
991 CTAAATGACTGTTTGAGACACGCTGTGACTGTGCCCTGCCAGTAATCTC 1040
456 .....PrometAspLysTyrLeuHisArgHisTrp 466
1041 AAATGACACAGTTCCTCCACAGCTGACAGTACAT.....CCCCGTGGG 1087
466 al.....LeuValTrpLeuAlaCysLeuLeuAlaAlaAlaLeuPhe 480
1088 TGTATGCTCTCATCACATCATCATCATCTGTGTGGGATGCTGTATC 1137
481 PhePheLeu.....LeuLeuLysLysAspArgArgLys 491
1138 GTGCGATCATCATGTATGACGTGAGGCTTTTGGCCCATCATCAAGAGA 1187
491 salAlaAlaArgLysSer..... 496
1188 ACATGCTGATGACTCCAAATCAATGATGCTTGGCCGTAGACAGACCTGA 1237
497 .....ArgThrAlaLeuLeuHisSerAlaAsp 506
1238 CTCGCCACACCTTGAGGCCGAGAGGTGTGATGCTGTACTGCGCCGAC 1287
507 GAlaAlaGlyTyrGluArgLeuValGlyAlaAlaLeuAlaSerAlaLeu...Se 522
1288 CACCCCTCTATGTGAGAGTGTGCTTAAGTTGCCGACGTTCCGTGATCAC 1337

```

```

522 rgLmetProLeuArgValAlaValAspLeuTrpSerArgArgGluLeuS 539
1338 TGCCTGTGACATGAGTACGCTTGCACCTGACCTGAGAGCAGTTATCT 1387
539 erAlaHisGlyAlaLeuAlaTrpPheHisHisGlnArgArgArgLysLeu 555
1388 CTGAGGTGGGGTCATGACCTGTGGTACCCGACAGAGAGAGAGATGGT 1437
556 GlnGluGlyGlyValValLeuLeuPheSer..... 566
1438 GAGAGCAACTCCAAATCATCATCTGTGTGCCAGGACCCCAAGCAAA 1487
567 .....ProAlaAlaValAlaGlnCysG 574
1488 GTGGAAGCTATCTTGGTGGTGGCTGAGCTGCTGCTCCAGTACGCTGTG 1537
574 lngInTrpLeuGlnLeuGlnThrValGluProGlyProHisAspAlaLeu 590
1538 ACCACTGG.....AAGCTGCT...GGGACCTTTTC 1566
591 AlaAlaTrpLeuSerCysValLeuProAspPheLeuGlnGlyArgAlaThr 607
1567 ACTGCACCCATGAGACATGATCTCTGCCAGACTTCAAGAGGCCGCTGCTT 1616
607 rgLysArgTyrValGlyValTyrPheAspGlyLeuLeuHisProAspSerV 624
1617 CGGCACCTAGCTGTGTGCTTACTTCTGATGCTGATGAGAGAGGATG 1666
624 alProSerProPheArgValAlaProLeuPheSerLeuPro..ThrGlnLe 640
1667 TCCCGACCTTTCACATCACTCCAGTACCCACTCATGAGCAAGATT 1716
640 uproAlaPheLeuAspAlaLeuGln.....GlyG 650
1717 GAGAGGTTTACTTCCGATCCAGACCTCGAGATGTTGAACCCGCCG 1766
650 LysSer.....ThrSerAlaGlyArgProAlaAspArgValGlu 663
1767 GATGACACCATGTCAGAGCTCAGAGGACAAATTAACCTGCAGAGCCCTTA 1816
664 ArgValThrGlnAlaLeuArgSerAlaLeuAspSerCysThrSerSerSe 680
1817 GTGCGCGGACGCTCAAGAGGCTGTGCTTGAAGTCCAGAGAGGCAACC 1866
680 rgLAlaPro 683
1867 CAGTCCCCG 1876

```

seq_name: /cgn2_6/ptodata/2/lna/58_COMB.seq:US-09-022-257-1

seq_documentation_block:

: Sequence 1, Application US/09022257

: Patent No. 6197525

: GENERAL INFORMATION:

: APPLICANT: Yao, Zhengbin

: APPLICANT: Spriggs, Melanie

: APPLICANT: Fanslow, William

: TITLE OF INVENTION: No. 6197525el Resptor: That Binds IL-17

: NUMBER OF SEQUENCES: 10

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Immunex Corporation

: STREET: 51 University Street

: CITY: Seattle

: STATE: WA

: COUNTRY: USA

: ZIP: 98101

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: Apple Power Macintosh

: OPERATING SYSTEM: Apple Operating System 7.5.5

: SOFTWARE: Microsoft Word for Apple, Version 6.0.1

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/09/022,257

FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/620,694
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USSN 08/410,535
 FILING DATE: 23 MARCH 1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia Anne
 REGISTRATION NUMBER: 34,695
 REFERENCE/DOCKET NUMBER: 2617-B
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206)587-0430
 TELEFAX: (206)
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3288 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Mouse
 STRAIN: HVS13 receptor
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 121..2715
 US-09-022-257-1
 Alignment_scores:
 Quality: 184.50 Length: 655
 Ratio: 0.666 Gaps: 32
 Percent Similarity: 42.290 Percent Identity: 20.763
 Alignment_block:
 US-09-899-471-5 x US-09-022-257-1 ..
 Align seg 1/1 to: US-09-022-257-1 from: 1 to: 3288

278SerGluArg. 280
 444 CCTGAGGGTGCAGAGCTGCTGCTGCAGCTGAAACCAATGAGCGGC 493
 281ValGluPheCysProPheArgGluAspProGlyAlaHisArgAsn 295
 494 TGTGTGTCAAGTTC...CAGTTTCTGTCCATGTGTCAGACATCCCGTAAG 540
 296 LeuTrpHisIleAlaArgLeuArg...ValLeuSerProGlyValTrpGly 311
 541 CGGTGGCGGTTTCTCCTCAGCCACTTGTGTAGATCCTGCG... 582
 311 nLeuAspAlaProCysCysLeuProGlyLysValThrLeuCysTrpGlna 328
 583CAGAGTATGAAGTGAAGTGT... 603
 328 laProAspGlnSerProCysGlnProLeuValProProValProGlnLys 344
 604CACCACCTGCCGGAAGCCCATCTCCTGATGCG 633
 345 Asn..AlaThrValAsnGluProGlnAspPheGlnLeuValAlaGlyHisP 361
 634 GACCCAAACCAACAAATCCA.....AGAT 656
 361 roAsnLeuCysValGlnValSerThr..... 369
 657 CATCTTGTGCTGACTGAGTGAGACAGACAGATGATGATGATGATGAT 706
 370TrpGluLysVa 373
 707 GCGTGAAGCTCAGGACGCTTTGGATCCCAACATCAGTGTGAGACCTGTG 756
 373 lGlnLeuGlnAlaCysSerTrpAlaAspSerLeuGlyProPheLysAsp 390
 757 GACACACAGCATCTGCGAGTGG..... 778
 390 spMetLeuLeuValGluMetLysThrGlyLeuAsnThrSerValCys 406
 779ACTTACCCCTGT 790
 407 AlaLeuGluPro.....SerGlyCysThrProLeu..ProSerMe 419
 791 GGAATGAAATCCACCCCTTACACAGGTCCTCTGGAAGATTTCTCTGACATCA 840
 419 tAlaSerThrArgAlaAlaArgLeuGlyGluLeuLeuGlnAspPhe 436
 841 GAGAACCAAGCTGCTTGTGATGTCGTTAAACAAATATTGTGGCCAGGCA 890
 436 rgSerHisGlnCysMetGlnLeuTrpAsnAspAspAspMetGlySerLeu 452
 891 AGAAGAAATTCATCAGCAGAGCTAATGACATTCACCTTAAGCAAGTTTC 940
 453 ..TrpAlaCys..... 455
 941 ACTGTGTGTCATACACAGTCGAGTCCAGCCCTTTCTTACAGAGCTGC 990
 455 455
 991 CTAATGACTGTTTGAGACAGCTGTGACTGTGCTGCCAGCAATATC 1040
 456PrometAspLysTrpIleHisTrpArgTrpV 466
 1041 AAATACCAAGTTCACCAAGCAGATTGACAGCTACAT...CCCTGTGGG 1087
 466 al.....LeuValTrpLeuAlaCysLeuLeuLeuAlaAlaLeuPhe 480
 1088 TGTATGGCTTCAATCAACATCATGATGATGATGATGATGATGATGATGAT 1137
 481 PhePheLeu.....LeuLeuLysLysAspArgArg 491
 1138 GTGCTGATCATCTGTATGACCTGAGAGCTTTTGTGGCCGATCAAGAGAA 1187
 491 sAlaAlaArgGlySer..... 496


```

1188 ACATGATGATGATCTCCAAATCATGCGATCTTCCCGTAGCAGACCTGA 1237
497 .....ArgThrAlaLeuLeuHisSerAlaAsp 506
1238 CTCCGCCACCCCTGAGGCCAGGAGGTGTGATCTTACTTGGCCGAC 1287
507 G1yAlaG1yTyrG1uArgLeuVal1yAlaLeuAlaSerAlaLeu...Se 522
1288 CACCCCTCTATGTGGAGGTGTCTTAAAGTTCGCCAGTCTCTGATCAC 1337
522 rG1uMetProLeuArgValAlaValAspLeuTrpSerArgArgG1uLeu 539
1338 TGCCGTGGGCACTGAGATAGCCCTTGACCTCTGTGAGAGCAGCTTATTC 1387
539 eR1aH1sG1yAlaLeuAlaTrpPheH1sH1sG1uArgArgG1eLeu 555
1388 CTGAGGTGGGGTCAATGATGACCTGGGTGAGCCGACAGACGAGATGTG 1437
556 G1uG1uG1yG1yVal1leLeuLeuPheSer..... 566
1438 GAGAGCAATCTCCAAATCATGATCTTCTTCCGAGCAGCAGCCAGCAAA 1487
567 .....ProAlaAlaValAlaG1uG1y 574
1488 GTGGAAGCATCTTGGGTGGCTGAGCTGCTGTCCAGCTACAGGTGTG 1537
574 InG1uTrpLeuG1uLeuG1uTrpValG1uProG1yProH1sAspAlaLeu 590
1538 AACACTGG.....AGCCTGCT...GGGACCTTTTC 1566
591 AlaAlaTrpLeuSerCysValLeuProAspPheLeuG1uG1yArgAlaTh 607
1567 ACATGACCATGAGATGATGCTGCTGACACTTCAAGAGCCAGCCTGCTT 1616
607 rG1yArgTyrValG1yVal1yTrpPheAspG1yLeuLeuH1sProAspSer 624
1617 CGGCACCTACCTGTTTGTCTGATGAGGCAATCTGTGTAGTGTGAGAGG 1666
624 alProSerProPheArgValAlaProLeuPheSerLeuPro...ThrG1uLe 640
1667 TCCCGACACCTTCAACATCACTCCAGGTACCCACTCATGACACATTT 1716
640 uProAlaPheLeuAspAlaLeuG1u.....G1yG 650
1717 GAGGAGGTTTACTTCCGATCCAGAGCTGAGATGTTTGAACCCGCGC 1766
650 LysSer.....ThSerAlaG1yArgProAlaAspArgValG1u 663
1767 GATGACACCATGTCAGAGGCTCACAGGCAATTAACCTGACAGACCTTA 1816
664 ArgValThrG1uAlaLeuArgSerAlaLeuAspSerCysThrSerSer 680
1817 GTGGCCGCGACCTCAAGCAGGCTGTGCTTGAAGTTCAGAGGTGGCAACC 1866
680 rG1uAlaPro 683
1867 CAGTGGCCCG 1876

```

seq_name: /cgn2.6/ptodata/2/lna/6A_COMB.seq:US-09-188-930-6

```

seq_documentation_block:
: Sequence 6, Application US/09188930A
: Patent No. 6150502
: GENERAL INFORMATION:
: APPLICANT: Watson, James D.
: APPLICANT: Strachan, Lorna
: APPLICANT: Sleeman, Matthew
: APPLICANT: Ornust, Rene
: APPLICANT: Morison, James Greg
: TITLE OF INVENTION: Compositions Isolated From Skin Cells
: FILE REFERENCE: 11000.1011c1

```

```

: CURRENT APPLICATION NUMBER: US/09/188,930A
: CURRENT FILING DATE: 1998-11-09
: NUMBER OF SEQ ID NOS: 348
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 6
: LENGTH: 1059
: TYPE: DNA
: ORGANISM: Mouse
: US-09-188-930-6

alignment_scores:
  Quality: 152.00      Length: 305
  Ratio: 1.094        Gaps: 13
  Percent Similarity: 45.574  Percent Identity: 23.607

alignment_block:
US-09-899-471-5 x US-09-188-930-6 ..

```

Align seg 1/1 to: US-09-188-930-6 from: 1 to: 1059

```

99 ProG1uG1uAlaG1ySerAspSerG1uLeuG1uSerArgAsn 115
115 LaserLeuAlaG1uValValLeuSerPheG1uAlaTrpProLeuAla 131
362 TTTCATCCCTCCATGACCATC...TCCACAGAGGCCAGCAGCAAA 408
132 ArgCysAlaLeuLeuG1uValG1uValProAlaAspLeuValG1uProG1 148
409 AGCGCCACCTTCACCTGCAGAGAGAGAGACATCTCCCTGAGACAG 458
148 Y...G1uSerValG1ySerAlaValPheAspCysPheG1uAlaSerLeu 164
459 GTACAAAAGTGTGAGAGCTGAA.....TTCTCTTTGATTTGC 499
164 L1yAlaG1uValG1uLeuTrpSerTyrThrLysProArgTyrG1uLysG1u 180
500 TCCCGAGGTGCGAGGCTTCCGGGTGACATCTCTCGAGGCCAGCCAGCC 549
181 LeuAsnLeuThrG1uLeuProAspCysArgG1yLeuG1uValArgAs 197
549 ..... 549
197 pSerLeuG1uSerCys.....TrpValLeuProTrpLeuAsnValSer 212
550 AGTGTGCGCTTTGTTATCAGTGGCCACTGGAATGTGAGACCTTGACTA 598
212 hTrAspG1yAspAsnValLeuLeuThrLeuAspValSerG1uG1uLys 228
599 GCCCTTTTATAC..... 612
229 PheSerPheLeuLeuTyrLeuArgProValProAspAlaLeuLysSer 245
612 ..... 612
245 uTrpTyrLysAsnLeuThrG1yProG1uAsnL1ySerLeuAsnH1sTyr 262
613 ...CAGAAAATTTGTCTGGAGGCCACACTGTGATGCTTATGAT 658
262 splLeuValProCysLeuCysL1eG1uValTrpSerLeuG1uProAspSer 278
659 TCCCTCTGCTGCTGATGATAGAGGCTTCCTACCTGACAGAGACT 708
279 G1uArgValG1uPheCysProPheArgLysAspProG1yAlaH1s...Ar 294
709 GTGAGCGCCAAAAGTGTCTCCAGAGCTGGCTGAGAGTATGTGCTC 758
294 GasLeuTrpH1sL1eAlaArgLeuArgValLeuSerProG1y..... 308
759 AGACTTCTGCAATACGCTTCACTGACTACAGCAGCAGCAATCAGA 808

```


309 ..ValTrpGlnLeuAspAlaProCysCysLeuProGlyIleValThrLeu 324
 809 TGGTGATGCTCTGACACTCCGCTCCACAGTAACTGGAGGCTCCCTC 858
 325 CysTrpGlnAlaProAspGlnSerProCysGlnProLeuValProProva 341
 859 TGTGTGGAGGAGGACCACTCAGACCCCTGCGAACCCTT..... 897
 341 lProGlnIleAsnAlaThrValAsnGlnProGlnAspPheGlnLeuValA 358
 898CCCAACGCCACAGCAGAGAGTCAAGAGATGTAATATCTTG 940
 358 lAgly.....HisProAsnIleuCysValGlnValSerThrTrpGlu 371
 941 AGAATGTGGACTTGACCCCGAGCTCTTAAAGTCTCA...TTTGA 987
 372 LysValGlnLeuGlnAlaCys.....SerTr 380
 988 AACAGCAGCCACGCTTGAATGTCCACACAGAGTGGCTCTCCATCTG 1037
 380 PalaAspSerLeu 384
 1038 GACTGTGAGCATG 1050

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-620-694A-9

seq_documentation_block:
 Sequence 9, Application US/08620694A
 Patent No. 5869286
 GENERAL INFORMATION:
 APPLICANT: Yao, Zhengbin
 APPLICANT: Spriggs, Melanie
 APPLICANT: Fanslow, William
 TITLE OF INVENTION: No. 5869286el Receptor That Blinds IL-17
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Immunex Corporation
 STREET: 51 University Street
 CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZIP: 98101
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: Apple Power Macintosh
 SOFTWARE: Microsoft Word for Apple, Version 6.0.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/620,694A
 FILING DATE: 21 MARCH 1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USSN 08/538,765
 FILING DATE: 7 AUGUST 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USSN 08/410,535
 FILING DATE: 23 MARCH 1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia Anne
 REGISTRATION NUMBER: 34,695
 REFERENCE/DOCKET NUMBER: 2617-B
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206)587-0430
 TELEFAX: (206)
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3223 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO
 ? ANTI-SENSE: NO
 ? ORIGINAL SOURCE:
 ? ORGANISM: Human
 ? STRAIN: IL-17 R (hCTLA8 receptor)
 ? NAME/KEY: CDS
 ? LOCATION: 93..2693
 ? US-08-620-694A-9

alignment_scores:
 Quality: 151.50 Length: 703
 Ratio: 0.493 Gaps: 37
 Percent Similarity: 43.670 Percent Identity: 21.479

alignment_block:
 US-09-899-471-5 x US-08-620-694A-9 ..

Align seg 1/1 to: US-08-620-694A-9 from: 1 to: 3223

55 SerAlaProGlyPro.....Valle 61
 120 GGTGTCCGGGGCCCTGCTGGGGCTGCTGCTGCTGCTGGGGCTGCT 169
 61 uValPro.....ThrArgLeuGlnThrGlnLeuValLeuArg 74
 170 GGGCCGGGTGGCGCCCTCCCTGGAGCTCTGGACACCGGGCGCTGCT 219
 74 ySPGlnIleuThrAspCysAlaLeuArgValArgValValHisLeu 90
 220 GCTCCAGCCG.....GGGCTAAACAGCAGCGTCAAGATATATATACC 260
 91 AlAlaValHisGlyHisTrpAlaGlnProGlnGlnAlaGlyLysSerAsp 107
 261 TGCCTGGATGACAGCTGGATTCACCTCA..... 290
 107 rGluLeuGlnIleuSerArgAsnAlaSerLeuGlnAlaGlnValLeu 124
 291 AACCTGACCCCTCTCTCCCAAGAGACCTGCAGATCCAGCTGACTT 339
 124 erPhe.....GlnAlaTrpProIleAlaArgCysAlaLeu 136
 340 CCCACACCCACACAGAGACCTGTTCCCGCTGCTC.....ATC 380
 137 GluValGlnValProAlaAspLeuValGlnProGlyLysSerValGly 153
 381 GAATGACACTCAGACAGAC..... 401
 153 lAlaValPheAspCysPheGlnAlaSerLeu.....GlyAlaG 166
 402GCCAGATCCCTGTTACCTGAGGAGGCTGAG 429
 166 lValAlaGlnIleTrpSerTrpThrLysProArgTrpGlnLysGlnLeu 182
 430 AGTATATCTGTC.....CTGCAGCTGATC 452
 183 LeuThrGlnIleuProAspCysArgGlyLeuGlnAlaGlyAspSer 199
 453 ACCAATGAACTTTG.....TGCCTCAGGTTGAGTTCTGCTCAACT 496
 199 eGln.....SerCysTrpValLeuProTrpLeuAsnValSerThr 213
 497 GAGGATCACCACAGCGGCTGCTTACCTCAGCCACTTTGTGTG 546
 213 SPGLysAsnValLeuLeuThrLeuAspValSerGlnGlnAspPhe 229
 547 ACCCTGAC.....CAGGATATATGAGGTG 569
 230 SerPheLeuLeuTrpLeuArgProValProAspAlaLeuLysSerLeu 246
 570 ACCGTACACCACTGCCAAGCCCATCTGATGGGACCAACCAACCA 619

430	AGTTATCTGTC.....	CTGCAGCTGCAC	452
183	LeuThhGlnGlnIleuProAspCysArgGlyLeuGluValAlaGaspSer11	199	
453	ACCAATGAAGAGTTTG.....TCGTCAGAGTTGAGTTCTGTGCCAAAT	496	
199	egln.....SerCysTrpValIleuProTryptLeuAsnValSerTha	213	
497	GAGGACATCACACAGCGGGGGCGGTTTTACCTTCACCCATTGTGGTGG	546	
213	spGlyAspAsnValIleuLeuThhLeuAspValSerGlnGluGlnAspHe	229	
547	ACCCGTAC.....CAGCAATTTGAGGTG	569	
230	SerPheLeuLeuTyrLeuArgProValProAspAlaLeuLysSerLeuTr	246	
570	ACCGTTACACCACTCGCCAAAGCCATCCCTGATGGGGCAACCAACACA	619	
246	PrpLysAsnLeuThhGlyProGlnAsnIleThhLeuAsnHisThAspL	263	
620	GTCCAGAAATTTG.....	632	
263	eulValProCysLeuCysIleGlnValTyrSerLeuGluProAspSerGlu	279	
633CTTGTGCTTGAATCTGTGAG	650	
280	ArgValGluPheCysProPheArgGluAspProGlyAlaHisArgAsnIle	296	
651	CAC.....	653	
296	uTrpHisIleAlaArgLeuArgValLeuSerPro.....G	308	
654GCCAGATGAAGATGATACCAAGCCATGATGAGCTCAGCA	693	
308	LYValTyrPglIleuAspAlaProCysCysLeuProGlyLysValThIle	324	
694	GCGTGTGGAGCCCAATCATCAGCG.....TGGAGACCCGTGAGGCCACAG	740	
324	uCysTrpGlnAlaProAspGlnSerProCysGln.....ProLeu	337	
741	CGCGGTGAGCT.....TCACCGTGTGAAGATCATCCCATTA	781	
338	ValProProValProGlnIleLysAsnAlaThrValAsnGluProGlnAspPh	354	
782	CCAGATCTCGTCGACGACCAACCATACCTGGCGCCACAG.....GGAGAACCAAG	823	
354	egIleuValAlaGlnHisProAsnLeuCysValGlnValSerThrTrpG	371	
824	TTGCTTTGAGCACATGCACCAATACCTGGCGCCACAG.....	861	
371	IuLysValGlnIleuGlnAlaCysSerTrpAlaAspSerLeuGlyProHe	387	
862CAGAAAGATTCCACACAGCATGCA	885	
388	LysAspAspMetLeuLeuValGlnMetLysThhGlyLeuAsnAsnThSe	404	
886	ACGTCACACTCATCTACGAAACCTTAAGGTGGTTC.....GCCAC	929	
404	rValCysAlaLeuIleuProSerGlyCysThrProLeuProSerMetAlas	421	
930	CAGATGCGATCCAGGCCCTTCTCAGCAGCT.....GCCCTCAATGACTG	973	
421	erThrArgAlaAlaIleuArgLysGluGluIleuLeuGlnAspHeArgSer	437	
974	CTCTAGACACTCCCGGAGCTG.....TTTCTGCTGC	1002	
438	HisGlnCysMetGlnLeuTrpAsnAspAspAsnMetLysLeuTrp	453	
1003	CAGAAATGCCAGACATGCCAAGCAATTCGGGAGCTACATGCCCTGTGG	1052	
454	AlaCysProMetAspLysTyrIleHisArgAlaTrpValLeuValTyrLe	470	


```

1053 GTC.....TACGTCTCATCAGGGCAT 1075
470 uAlaCysLeuLeuLeuAlaAlaLeuPhePheLeuLeuLeuAla 487
      |||||.....
1076 CTCATCTCTCTGCTGGCTCCGTCATC.....CTGCTCATCTGTGCA 1119
487 yAspArgArgLysAlaAlaArgLysSer..... 496
      |||||.....
1120 TGACCTGAGGCTAGCTGGCTGGAAAGTGAATAATACATGATGACACC 1169
496 ..... 496
1170 AATACACCATGAGCTGCTGGCTGACCTGATCCGCCACCGCTGA 1219
497 ....ArgThAlaLeuLeuLeuHisSerAlaAspLysAlaGlyTyrGlu 512
      |||||.....
1220 GCCCAGGAGAGCTGTGATCATCTACTACGACGACACCCCTTAAGTGG 1269
512 rGLeuValGlyAlaLeuAlaSer...AlaLeuSerGlnMetProLeuArg 527
      |||||.....
1270 ACCTGTCTCTGAATTCGCCAGTCTCTCACCGCTGCGCAGCGAA 1319
528 ValAlaValAlaPheLeuTyrPserArgArgLysLeuSerAlaHisGlyAlaLe 544
      |||||.....
1320 GTGGCCCTGACCTGCTGGAAGACAGCCCATCTCGAGCGAGAGTCAT 1369
544 uAlaTyrPheHisGlnArgArgArgLysLeuGlnGlnGlyGlyVal 561
      |||||.....
1370 GACCTGGGTGGCGCTCAGAAAGCAGAGATGTGGAGAACACTTAANA 1419
561 AlLeuLeuPheSerProAlaAlaValAlaGlnCysGlnGlnTyrLeu 577
      |||||.....
1420 TCATCTCTCTGCTCCCGCGCAGCGCCAAAGTGGCAGCGCTCCG 1469
578 .....GlnLeuGlnThrValGlnProGlyPro...His 587
1470 GCGCGGGGGCGCTGTGCGGCTGCGTCCGACCGAAGACCCGCTGG 1519
587 sAspAlaLeuAlaAlaTyrPheSerCysValLeuProAspPheLeuGln 604
      |||||.....
1520 GGACCTGTTCACCTGACGACCATGACATGATCTCCGAGCTCAAGAGGC 1569
604 LArgAlaThrGlyArgTyrValGlyTyrPheAspLysLeuLeuHis 620
      |||||.....
1570 CAGCCCTGCTGGCACCCTACTCTGCTCTTCCAGAGAGTCACTGT 1619
621 ProAspSerValProSerProPheArgValAlaProLeuPheSerLeuPr 637
      |||||.....
1620 GACGGGAGCTCCCGACCTGTTGCGCGCGCGCGGTACCCGCTC.. 1667
637 oThrGlnLeuProAlaPheLeuAspAlaLeuGlnGlnLysSerThr 654
1667 ..... 1667
654 eAlaGlyArgProAlaAspArgValGluArgValThrGlnAlaLeuArg 670
1668 .....ATGACAGGCTTCGAGAGTGTCTCCGCTCCGCTCAG 1703
670 gSerAlaLeuAspSerCysThrSerSerSerGlnAlaProGlyCysCys 686
      |||||.....
1704 GACCTGGAATGTTCCAGCGCGCGCGCATGACCGCGTGGAGAGCTGT 1752
      |||||.....
Seq. name: /cgn2_6/ptodata/2/lna/6A.COMB.seq:us-09-022-696-9
Seq. documentation block:
: Sequence 9, Application us/09022696
: Patent No. 6072037
: GENERAL INFORMATION:
: APPLICANT: Yao, Zhengbin
: APPLICANT: Spriggs, Melanie
: APPLICANT: Fanslow, William
: TITLE OF INVENTION: No. 6072037el Receptor That Binds IL-17
: NUMBER OF SEQUENCES: 10

```

```

CORRESPONDENCE ADDRESS:
: ADDRESSEE: Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: Apple Power Macintosh
: OPERATING SYSTEM: Apple Operating System 7.5.5
: SOFTWARE: Microsoft Word for Apple, Version 6.0.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/022,696
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/620,694
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: USSN 08/410,535
: FILING DATE: 23 MARCH 1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Perkins, Patricia Anne
: REGISTRATION NUMBER: 34,695
: REFERENCE/DOCKET NUMBER: 2617-B
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206)587-0430
: TELEFAX: (206)
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3223 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Human
: STRAIN: IL-17 R (hCTLA8 receptor)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 93...2693
: US-09-022-696-9

alignment_scores:
: Quality: 151.50 Length: 703
: Ratio: 0.493 Gaps: 37
: Percent Similarity: 43.670 Percent Identity: 21.479

alignment_block:
: US-09-899-471-5 x US-09-022-696-9
Align seg 1/1 to: US-09-022-696-9 from: 93 to: 3223
55 SerAlaProGlyPro.....ValLe 61
      |||||.....
120 GCTGTCCCGGGGCCCTGCTGGGCTGCTCTGCTCTGCTGCGCGTCT 169
61 uValPro.....ThrArgLeuGlnThrGlnLeuValLeuArg 74
      |||||.....
170 GCGCCCGGTGGCGCTCCCTGCTGCTGACACCGCGCGCTGCTGT 219
74 ySProGlnLysThrAspCysAlaLeuArgValArgValValHisLeu 90
      |||||.....
220 GCTCCACGCG.....GGCTTAATCTGCACGGTCAAGATATGATACC 260
91 AlaValHisGlyHisTyrAlaGlnProGlnGlnAlaGlyLysSerAspSe 107
      |||||.....

```


388 LysAspArmLeuValGluMetLysThrGlyLeuAsnThrSe 404
 886 ACGTACACTCACTATCACCACTTAAAGGTCCTGC.....GCCAC 929
 404 rValGlyAlaLeuGluProSerGlyCysThrProLeuProSerMetLas 421
 930 CAAATGCAGATCCAGCCCTTCTTCACAGCT.....GCCCAATGACTG 973
 421 eThrArgAlaAlaArgLeuGluGluLeuLeuGlnAspPheArgSer 437
 974 CCTCAGACACTCCGGAGCTG.....TTTCCTGCC 1002
 438 HisGlnCysMetGlnLeuThrAsnAspAspMetGlySerLeuTrp 453
 1003 CAGAAATGCCACACACTCCAGAACCAATTCCGAGCTACATGCCCTGTGG 1052
 454 AlaCysProMetAspLysTrpLleHisArgTrpValLeuValTrpLe 470
 1053 GTG.....TACTGTTCATCAGGGGCAT 1075
 470 uAlaCysLeuLeuAlaAlaAlaLeuPhePhePheLeuLeuLysL 487
 1076 CTCATCTCTGCTGGTGGGCTCCGTATC.....CTCTCATCTGTGCA 1119
 487 yAspArgArgLysAlaAlaArgLysSer..... 496
 1120 TGACCTGAGGCTACTGTGGCCTGGAAAGTGAATAATACACTGATGACACC 1165
 496 496
 1170 AAATACACCGATGCCCTGCTGCGCTGACTGATCCCCACCGCTGAA 1219
 497 ...ArgThrAlaLeuLeuHisSerLysArgLysAlaGlyTrpGlu 512
 1220 GCCACGAGAGCTGTGATCATGTACTCAGCCGACACACCCCTCTATGTCG 1269
 512 rGluValGlyAlaLeuAlaSer...AlaLeuSerGlnMetProLeuArg 527
 1270 ACGTGGCTCTGAATATCGCCAGTTCCTGCTACACCGCTCCGGCAGCGAA 1319
 528 ValAlaValAspLeuTrpSerArgArgGluLeuSerAlaHisGlyAlaLe 544
 1320 GTGGCCCTCGAGCTGCGAAGAGCGGCCATCTCGAGGACAGAGATCAT 1369
 544 uAlaTrpPheHisGlnArgArgGluLeuGlnGluGlyValVal 561
 1370 GACCTGGGTGGCGCTCAACACAGAGATGTGGAGAGCACTGTAA 1419
 561 alIleLeuLeuPheSerProAlaAlaValAlaGlnGlyGlnGlnTrpLeu 577
 1420 TCATCTGCTGTGCTCCGGCGGACCGCGGCCAAGTGGCAGGGGCTCTG 1469
 578GlnLeuGlnTrpValGluPro...His 587
 1470 GGCCGGGGGGCGCTGTGCGGTGCGCTCGACACAGAAAGCCCGTGG 1519
 587 sAspAlaLeuAlaAlaTrpLeuSerCysValLeuAspArgPheLeuGln 604
 1520 GGACCTGTTCACTGACGCAATACATGATGTCCTCCCGACTTAAAGAGG 1569
 604 LysArgAlaThrGlyArgTrpValGlyValTrpPheArgGlyLeuHis 620
 1570 CAGCTGCTTGGGACACACAGTCTGCTACTTACGCAAGTCACTGT 1619
 621 ProAspSerValProSerProPheArgValAlaProLeuPheSerLeuTrp 637
 1620 GACGGCAGCGTCCCGACCTGTTCGGCGGGCGCGGTACCCGCTC... 1667
 637 oThrGlnLeuProAlaPheLeuAspAlaLeuGlnGlyLysCysSerThr 654
 1667 1667

654 eraIaGIaTgPrOaIaSpArGyValGlu.ArgValaThrGlnAlaLeuAr 670
1668ATGACACAGGTTCGAGAGGTGTACTTCCGATCCAG 1703
670 gSerAlaLeuAspSerCysThrSerSerSerAlaProGlyCysCys 686
1704 GACCTGGAGATGTTCCACCGCGCGCATGACACCGCTGAGGGCAGCTGT 1752

seq_name: /cgn2_6/plodata/2/lna/6A_COMB.seq:us-08-978-773-3

seq_documentation_block:

Sequence 3, Application US/08978773

Patent No. 6083906

GENERAL INFORMATION:

APPLICANT: Trout, Anthony

TITLE OF INVENTION: Method of Regulating Nitric Oxide Production

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple PowerMacintosh

OPERATING SYSTEM: Apple Operating System 7.5.5

SOFTWARE: Microsoft Word for PowerMacintosh, Version 6.0.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/978,773

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 60/052,525

FILING DATE: 27 NOVEMBER 1996

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,693

REFERENCE/DOCKET NUMBER: 2623-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)587-0430

TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 3223 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Human

IMMEDIATE SOURCE:

CLONE: IL-17R

FEATURE:

NAME/KEY: CDS

LOCATION: 93..2690

US-08-978-773-3

alignment_scores:

Quality: 151.50

Ratio: 0.493

Percent Similarity: 43.670

Percent Identity: 21.479

alignment_block:

US-09-899-471-5 x US-08-978-773-3

Align seg 1/1 to: US-08-978-773-3 from: 1 to: 3223

55 SerAlaProGlyPro.....Valle 61
120 GCGTGTCCGGGGCCCTGCTGGGGCTCTCTCTGCTCTGGGGCTCT 169
61 uValPro.....ThrArgLeuGlnThrGluLeuValLeuArgC 74
170 GGGCCCGGCGGGCGCTCCCTGCGATCTGAGACCGCGCTGCTCT 219
74 ySerProGlnyThrAspCysAlaLeuArgValArgValValHisLeu 90
220 GCTCCACGCG.....GGCTAACTGACAGGTCAGAAATAGTACC 260
91 AlaValHisGlyHisThrAlaGluProGluValGlyLysSerAspSe 107
261 TGCGTGGATGACAGCTGGATTCCACCTCGA..... 290
107 rGluLeuGlnGluSerArgAsnAlaSerLeuGlnAlaGlnValValLeu 124
291 .AACCTGACCCCTCCCTCCCAAAGACCTGACATCCAGCTGCACTTG 339
124 erPhe.....GlnAlaTyProIleAlaArgCysAlaLeuLeu 136
340 CCCACACCCACAAGAGACCTGTTCCCGCTGCTCAC.....ATC 380
137 GluValGlnValProAlaAspLeuValGlnProGlyInserValGlySe 153
381 GAATGACACCTGACAGACAGAC..... 401
153 rAlaValPheAspCysPheGlnAlaSerLeu.....GlyAlaG 166
402GCCACATCCCTGATCTCAGCTGAGTGACG 429
166 LuValGlnIleTyrSerTyrThrLysProArgTyrGlnGluLeuAsn 182
430 AGTATATCTGTC.....CTGCAGCTGAC 452
183 LeuThrGlnGlnLeuProAspCysArgGlyLeuGlnValArgAspSerI 199
453 ACCAATGAACGTTTG.....TGCGTCAAGTTTACGTTCTGTCACAACT 496
199 eGln.....SerCysTyrValLeuProTyrLeuAsnValSerThra 213
497 GAGGCATCACACACAGCGGCTGCTTACCTCAGCACTTGGGTG 546
213 spGlyAspAsnValLeuLeuThrLeuAspValSerGluGlnGlnAsphe 229
547 ACCCTGAC.....CAGGAATATGAGTG 569
230 SerPheLeuLeuTyrLeuArgProValProAspAlaLeuLysSerLeu 246
570 ACCGTCACACACCTGCCCAAGCCATCCCTGATGGGACCCCAACACCA 619
246 pTyrLysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAsp 263
620 GTCCAAGAAATTC..... 632
263 euValProCysLeuCysIleGlnValThrSerLeuGluProAspSerGlu 279
633CTTGCTGCTGACTGTAG 650
280 ArgValGluPheCysProPheArgGluAspProGlyAlaHisArgAsnLe 296
651 CAC..... 653
296 uTPrHisIleAlaArgLeuArgValLeuSerPro.....G 308
654GCCAGATGAAGTAAACACAGCCATGATGATGATGACCA 693
308 lValThrGlnLeuAspAlaProCysCysLeuProGlyLysValThrLe 324
694 GCCTGTGGAGACCCCAACATCACCG...TGGAGACCTGAGGCGCCACAG 740
324 uCysTyrGlnAlaProAspGlnSerProCys.Gln.....ProLeu 337

1. HYPOTHETICAL: NO
 2. ANTI-SENSE: NO
 3. ORIGINAL SOURCE:
 4. ORGANISM: Human
 5. STRAIN: IL-17 R (NCT1148 receptor)
 6. FEATURE:
 7. NAME/KEY: CDS
 8. LOCATION: 93..2693
 9. IS-09-022-253-9

[illegible]

alignment_block:
US-09-899-471-5 x US-09-022-253-9 .

Align seg 1/1 to: US-09-022-253-9 from: 1 to: 3223

55 SerAlaProGlyPro.....Valle 61
 120 GGTGTCCGGGGCCCTGCTGGGGCTCTCTGCTGCTCTGGGGCTGCT 169
 61 uValPro.....ThArgLeuGlnThrGluLeuValLeuArgC 74
 170 GGGCCCCGGGTGGCGCCCTCCCTGCGACCTCTGGACCACGGCGGTGGTCT 219
 74 ysProGlnIysThrAspCysAlaLeuArgValaArgValValaHisLeu 90
 220 GCTCCCAAGCG.....GGGCTAAACTGCACGGCTCAAGAATAGTACC 260
 91 AlaValHisGlyHisTrpAlaGluProGlnGluAlaGlyLysSerAspSe 1077
 261 TGCCTGGATGACAGCTGGATTCCACCTCGA..... 290
 107 rGluLeuGlnIleSerArgAsnAlaSerLeuGlnAlaGlnValValLeus 124
 291 .AACCTGACCCCTCCCTCCCAAGGACCTCAGATTCAGCTGACACTTGG 339
 124 erPhe.....GlnAlaLeuTrpProIleAlaArgCysAlaLeuLeu 136
 340 CCCACACCCACACAGAGACCTGTTCCTCCCGGGCTCAC.....ATC 380
 137 GluValGlnValProAlaAspLeuValGlnProGlnIleSerValGlySe 1533
 381 GAATGGACACTGCAGACAGAC..... 401
 153 rAlaValPheAspCysPheGlnAlaSerLeu.....GlyAlaG 166
 402GCCAGATCTCTGACTCGAGGCGAG 429
 166 IuValGlnIleTrpSerTyrThrIysProAlaGlyGlnLeuAsn 1822
 430 ACTTACTCTGC.....CTGACAGCTGAC 452
 183 LeuThrGlnGlnLeuProAspCysArgGlyLeuGlnValArgAspSerI 199
 453 ACCAATGAACGTTTG.....TGGCTCAGGTTTGAGTTCTGTGCCAACT 496
 199 eGln.....SerCysTrpValLeuProTrpLeuAsnValSerTha 213
 497 GAGGCAATCACACAGCGGTGGGCTTTACTCTTACCGCACCTTGTTGGTTG 546
 213 spGlyAspAsnValLeuLeuThrLeuAspValSerGlnGlnIleAspPhe 229
 547 ACCCTCAGC.....CAGGAATATGAGGTG 569
 230 SerPheLeuLeuTyrLeuArgProValProAspAlaLeuLysSerLeu 246
 570 ACCGTTTCACCACTGCCCAAGGCCCATCTCTATGAGGAGACCAACCAACCA 619

246 pTyLysAsnLeuThrGlyProGlnAsnIleThrLeuAsnHisThrAsp 263
620 GTCCAGAAATTC..... 632
263 euValProCysLeuCysIleGlnValTrpSerLeuGluProAspSerGlu 279
633 111 111111 111
280 ArgValGluProCysProPheArgGluAspProGlyAlaHisArgAsnLe 296
651 CAC..... 653
296 uTrpHisIleAlaArgLeuArgValLeuSerPro.....G 308
654GCCAGAGTGAAGGTATACCACCCATGCATGACGCTCAGCA 693
308 lValValTrpGlnLeuAspAlaProCysCysLeuProGlyLysValThrLe 324
694 GCCTGTGGGAGCCCCAACATCACC...TGAGAGACCTGTGAGGCCACAC 740
324 uCysTrpGlnAlaProAspGlnSerProCysGln.....ProLeu 337
741 CACGCGGTGAGCT.....TCACCTGTGGAGCAAGATCTACCATTA 781
338 ValProProValProGlnLysAsnAlaThrValAsnGluProGlnAsp 354
782 CCAGATCGTCGTCACCAAGTTTCCGCACAT...GGAGAACCCAG... 823
354 eGlnLeuValAlaGlyHisProAsnLeuCysValGlnValSerThrTrp 371
824 TTGCTTGACGACAGTCACCACTACTGCGCCACAG..... 861
371 lValValGlnLeuGlnAlaCysSerTrpAlaAspSerLeuGluProPhe 387
862CAGAGAGTTTCCACACGACATCA..... 885
388 LysAspAspMetLeuLeuValGluMetLysTrpGlyLeuAsnAsnThrSe 404
886 ACGTCACTACTCTTACGCAACCTTAAAGGCTGTC.....GCCAC 929
404 rValCysAlaLeuGluProSerGlyCysThrProLeuProSerMetAla 421
930 CAAGTGCAGATCCAGCCCTTTCAGCAGCT....GCCATATGACTG 973
421 eTrpTrpArgAlaAlaArgLeuGlyGluGluLeuLeuGlnAspPheArgSer 437
974 CCTCAACACTCGCGCAGCT.....TTTCCTGCC 1002
438 HisGlnCysMetGlnLeuTrpAsnAspAspAspMetGlySerLeuTrp 453
1003 CAGAAATGCCAGACACTCCAGAAACAAATCCGACTACATGCCCTGTGG 1052
454 AlaCysProMetAspLysLysIleHisArgArgTrpValLeuValTrpLe 470
1053 GTG.....TACTGGTTCATCCGGGCAT 1075
470 uAlaCysLeuLeuLeuAlaAlaAlaLeuPhePheLeuLeuLysL 487
1076 CTCACATCTGCTGTGGGCTCCGATC.....CTGCATCTGCTGCA 1119
487 yAspArgArgLysAlaAlaArgGlySer..... 496
1120 TGCACCTGAGGCTAAGTGGGCTGGAATCGAATAATACAGTATGACACC 1169
496 496
1170 AAATACACCGATGCGCTGCGGCTGACTGATCCCCACCGCTGA 1219
497 ...ArgThrAlaLeuLeuLeuHisSerAlaAspGlyAlaGlyTrpGlu 512
1220 GCCACAGGAAGGTCTGATCATCTACTGACGCCAGCACCCCTTACGTGG 1269
512 rGluValGlyAlaLeuAlaSer...AlaLeuSerGlnMetProLeuArg 527


```
1270 ACCTGGTCTGAAATTCGCCAGTTCCTGCTACCCGCGCCGACGGAA 1319
      ::|||:::||||:::
528 ValAlaValAspLeuTrpSerArgArgIleuSerAlaHisGlyAlaLe 544
      |||||:::|||||:::
1320 GTGGCCCTGCGACTGGAAGACAGACATCTCGAGGACAGAGTCAAT 1369
      ::|||:::|||||:::
544 uAlaTrpPheHisGlnArgArgIleLeuGlnGlyValVal 561
      ::|||:::|||||:::
1370 GACCTGGGTGGCGTCAGACGAGATGCTGGAGCACTCAAGA 1419
      ::|||:::|||||:::
561 AlIleLeuPheSerProAlaValAlaGlnCysGlnIleuTrpLeu 577
      ::|||:::|||||:::
1420 TCATGCTCTCTGCTCCCGCGGACGCGCCCAAGTGCAGGCGCTCTG 1469
      ::|||:::|||||:::
578 .....GlnLeuGlnThrValGluPro...His 587
      ::|||:::|||||:::
1470 GGCCGGGGGGCGCTGTGCGGCTGCGCACGGAAGCCCGTGGG 1519
      ::|||:::|||||:::
587 sAspAlaLeuAlaIleTrpLeuSerCysValLeuProAspPheLeuGln 604
      |||:::|||||:::
1520 GACCTGTTCTACTGCAACCATGAACATGATCTCCCGACTTCAAGAGGC 1569
      ::|||:::|||||:::
604 LysArgAlaThrGlyArgTyrValGlyValTyrPheAspGlyLeuLeuHis 620
      ::|||:::|||||:::
1570 CAGCGTCTTCGGCAGCTACGTAAGTCTCTCAGCAGGTCAGCTGT 1619
      ::|||:::|||||:::
621 ProAspSerValProSerProPheArgValAlaProLeuPheSerLeuPr 637
      ::|||:::|||||:::
1620 GACGGGCGAGTCCCGACGCTGTGCGGCGCGCGGTACCGGCTC... 1667
      ::|||:::|||||:::
637 oThrGlnLeuProAlaPheLeuAspAlaLeuGlnGlyCysSerThrS 654
      ::|||:::|||||:::
1667 ..... 1667
654 erAlaGlyArgProAlaAspArgValGluArgValThrGlnAlaLeuAr 670
      |||||:::|||||:::
1668 .....ATGACAGGTTTCAGAGAGGTGTACTTCGCAATCCAG 1703
      ::|||:::|||||:::
670 gSerAlaLeuAspSerCysThrSerSerGlnAlaProGlyCysCys 686
      ::|||:::|||||:::
1704 GACCTGAGATGTTCCAGCGCGGCGCATGCACCGCGTACGGGAGAGCTGT 1752
      ::|||:::|||||:::
```


CM of: US-09-899-471-5 to: EST: * out_format : pfs

Date: Sep 28, 2002 9:43 PM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODE=firmet -p2n.model -DEV=x1h
-O/cgn2.1/USPTO.spool/US09899471/funat.27092002.142439.17896/app-query.fasta.1.2519
-DB=EST -OEM=fastap -SUFFIX=rest -GAPOP=12.000 -GAPEXT=4.000
-MINMAPTH=0.100 -LOOPT=0.000 -LOOEXT=0.000 -GAPOP=4.500
-GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FAPEXT=7.000 -START=1 -MATRIX=blotsum62 -TRANS=human40.cdi
-LIST=45 -DOCMALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0
-ALIGN=15 -MODE=LOCAL -OUTMT=pfs -NORM=ext -HEADSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09899471@cgn1.1.7278
-CPU=6 -ICPU=3 -LONGLOG -DEV=TIMEOUT=120 -WARN.TIMEOUT=30
-NO_XLPRX -WAIT -THREADS=1

Search information block:

Query: US-09-899-471-5
Query length: 698
Database sequences: 13736207
Database length: -1841457050
Search time (sec): 7183.320000

Score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
est:AV010326	+ 1418.50	2361.66	2.7e-122	1039	AV010326 AV010326 Mus musculus
est:BB609618	+ 1303.50	2168.66	1.5e-111	1006	BB609618 BB609618 Riken full-1
est:BI332055	+ 1226.00	2042.42	1.6e-104	697	BI332055 602982878F1 NCI CGAP
est:BG688034	+ 1187.00	1975.54	8.6e-101	784	BG688034 602832421F1 NCI CGAP
est:AI007139	+ 1095.00	1820.98	9.2e-92	775	AI007139 ua81f02.1.1 Soares.mamm
est:BF237179	+ 1058.00	1759.56	4.7e-89	722	BF237179 602028133F1 NCI CGAP
est:BB625706	+ 1039.00	1728.90	9.7e-87	643	BB625706 BB625706 RIKEN full-1
est:BI251117	+ 1012.50	1681.39	4.1e-84	839	BI251117 602984315F1 NCI CGAP
est:BB637050	+ 1007.00	1675.50	4.4e-84	620	BB637050 BB637050 RIKEN full-1
est:BB290565	+ 999.00	1663.14	2.2e-83	562	BB290565 601089427F1 NCI CGAP
est:BF160122	+ 975.00	1621.43	4.6e-81	635	BF160122 601767822F1 NCI CGAP
est:BG619177	+ 912.50	1512.05	6.7e-75	935	BG619177 602178928F1 NCI CGAP
est:BG917539	+ 908.00	1506.40	1.2e-74	787	BG917539 602817086F1 NCI CGAP
est:BI765330	+ 902.00	1494.55	5.3e-74	922	BI765330 603050641F1 NIH_MGC
est:BI282632	+ 893.00	1483.84	2.1e-73	619	BI282632 UT-R-CW05-ced-b-11-0-U
est:BM210218	+ 876.00	1456.35	7.1e-72	561	BM210218 C0660H10-3 NIA Mouse T
est:BM009552	+ 872.50	1447.98	2.1e-71	702	BM009552 603630118F1 NIH_MGC
est:BI689749	+ 860.00	1425.11	3.9e-70	829	BI689749 603316255F1 NCI CGAP
est:AA238911	+ 852.00	1418.22	9.5e-70	459	AA238911 mx92804.1.1 Soares.mous
est:BM010282	+ 834.00	1387.48	4.9e-68	878	BM010282 2M0281A07F1 NIH_MGC
est:BI770219	+ 822.00	1360.88	1.5e-67	609	BI770219 603053287F1 NIH_MGC
est:BI820026	+ 814.50	1347.87	7.9e-66	885	BI820026 603037207F1 NIH_MGC
est:BF161932	+ 805.50	1336.13	3.6e-65	653	BF161932 601767930F1 NCI CGAP
est:BI913818	+ 797.50	1319.36	3.0e-64	880	BI913818 603180205F1 NIH_MGC
est:W13694	+ 791.00	1305.78	1.7e-63	1118	W13694 ms94902.01 Soares.mous
est:BI832635	+ 790.00	1308.71	1.2e-63	796	BI832635 603082012F1 NIH_MGC
est:BB760916	+ 778.00	1292.92	9.1e-63	497	BB760916 BB760916 RIKEN full-1
est:BI821485	+ 769.50	1272.42	1.3e-61	869	BI821485 603038380F1 NIH_MGC
est:BI767178	+ 765.50	1266.00	2.9e-61	846	BI767178 603060554F1 NIH_MGC
est:BI819336	+ 762.50	1260.74	5.6e-61	862	BI819336 603037792F1 NIH_MGC
est:BI753674	+ 754.00	1258.31	7.7e-61	855	BI753674 603032357F1 NIH_MGC
est:BB687616	+ 751.00	1248.24	2.8e-60	893	BB687616 602766939F1 NCI CGAP
est:BI770842	+ 753.50	1245.22	4.1e-60	893	BI770842 603058331F1 NIH_MGC
est:BI751043	+ 751.00	1246.89	3.3e-60	846	BI751043 602986759F1 NCI CGAP
est:AV752013	+ 749.00	1242.00	5.7e-60	567	AV752013 AV752013 NPJ Homosap
est:BI755515	+ 749.00	1242.70	5.7e-60	462	BI755515 209379 MRC 2Bov Bos t
est:AI463827	+ 728.50	1203.93	8.2e-58	835	AI463827 va32303.01 Quaymofor
est:BF101574	+ 725.50	1198.58	1.6e-57	858	BF101574 601733388F1 NCI CGAP
est:BI767700	+ 725.50	1198.58	1.6e-57	858	BI767700 603060565F1 NIH_MGC

seq_name: gb_estl:AV010326
gb_estl:BB736394 + 724.00 1203.38 8.8e-58 444 | BB736394 BB736394 RIKEN full
gb_estl:BI490389 - 724.00 1196.70 2.1e-57 810 | BI490389 603031886F1 NIH_MGC
gb_estl:BF791762 + 716.50 1182.37 1.3e-56 946 | BF791762 602251953F1 NIH_MGC
gb_estl:BI765023 - 715.00 1181.65 1.4e-56 804 | BI765023 603051160F1 NIH_MGC
gb_estl:AM045329 - 713.00 1185.03 9.3e-57 438 | AM045329 UT-M-BH1-aku-h-11-c

seq_name: gb_estl:AV010326

seq documentation block:
LOCUS AV010326 1039 bp mRNA linear EST 23-OCT-2001

DEFINITION AV010326 Mus musculus 18-day embryo C57BL/6J Mus musculus cDNA
clone 1110025H02, mRNA sequence.

ACCESSION AV010326
VERSION AV010326.2 GI:16356130

KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1039)

AUTHORS

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Kono, H., Kouda,
M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toy, T.,
Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

Unpublished (2001)
On May 11, 1999 this sequence version replaced gi:4787313.

CONTACT: Chile Owa
Genome Science Laboratory
RIKEN

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098

Email: genome-res@r.riken.go.jp
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Matahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura,
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multichannel sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa,
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.

Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Location/Qualifiers

1. 1039
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1110025H02"
/clone_lib="Mus musculus 18-day embryo C57BL/6J"


```

/sex="mixed"
/dev.stage="18-day embryo"
BASE COUNT      175 a      353 c      311 g      199 t      1 others
ORIGIN

Alignment_scores:
  Quality: 1418.50      Length: 343
  Ratio: 4.489      Gaps: 8
  Percent Similarity: 92.128      Percent Identity: 87.464

Alignment_block:
US-09-899-471-5 x AV010326 ..

Align seg 1/1 to: AV010326 from: 1 to: 1039

360 HSPProteinLeuValGlnValSerThrTrpGluLysValGlnLeuGlu 376
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
16 CACCCACACTCTCC.....TCCAGTGGAGAGCTGGAGAGGCTCAG 56

376 n.Ala..CysSerTrpAlaAspSerLeuGlyProPheLysAspMet 391
    :||| ||| ||| :|||:::|||||:::|||||:::|||||:::
57 GTCCAGCTCTCTGCTGCTG...ACTCTGGCCCACTTAAGATGATAT 102

392 LeuLeuValGlnMetLysThrGlyLeuAsnAsnThrSerValCysAla 408
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
103 CTGTAGTGAATGAACACCGCTCAACACCAATGATCTCTG...TT 149

408 uGluProSerGlyCysThrProLeuProSerMetValSer..ThrAla 424
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150 GGAACCAACGCTGC..TGTAACACCTGCGAGAGCTGCTCAACGAGAG 198

425 AlaArgLeuGlyGluGlnLeuGlnAspPheArgSerHisGlnCysMe 441
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
199 GCTGGCTGGAGAGAGAGCTGCTCAAGACCTCCATCAGACAGGTAT 248

441 tGlnLeuTrpAsnAspAspMetGly..SerLeuTrpAlaCysPromet 457
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
249 GCAAGCTGGAACGATGACAAATGATCGCTATG...GCTGCCCCATG 295

458 AspLysTrpTrpLeuAspArgTrpValLeuValTrpLeuAlaCysLeu 474
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
296 GACAACTACTCCACAGCGCTGGCTCTATATGGCTGGCTGCTACT 345

474 uLeuAlaAlaAlaLeuPhePheLeuLeuLysLysAspArgL 491
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
346 CTGGCTGGCGCGCTTTTCTCTCCCTCTCTAAAGAGCCGAGAGA 395

491 yAlaAlaAlaArgGlySerArgThrAlaLeuLeuHisSerAlaAspGly 507
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
396 AAGCGGCGCTGGCTCCCGACGCGCTTGCCTCCACTCCGCGAGCGA 445

508 AlaGlyTrpGluArgLeuValGlnAlaLeuAlaSerAlaSerGlnMe 524
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
446 GCGGCTAGCAGAGCTCTGCTGGAGACACTGGCGTCCGGTGAAGCAG 495

524 tProLeuArgValAlaValAspLeuTrpSerArgArgGlnLeuSerAla 541
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
496 GGCACATGCGCGTGGCGCTGGACCTGGAGCGCGCGAGCTGAGACGCC 545

541 tAGlyAlaLeuAlaTrpPheHisGlnArgArgGlnLeuGlnGlu 557
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
546 ACGGAGCCCTAGCCTG..TTCACACACAGCAGCGCCGATCTCGAGAG 594

558 GLyGlyValAlaIleLeuLeuPheSerProAlaAlaValAlaGlnCys 574
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
596 GGTGGCGTGGATCTCTCTCTCTCCCGCGCGCGCGGCGAGCTGTA 644

574 nGlnTrpLeuGlnLeuGlnThrValGlnProGlyProHisAspAlaLeu 591
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
645 GCAAGTGGCTGACAGCTCCAGACAGTGGAGCCCGCGCATGACGCCCTCG 694

592 tAlaLeuTrpLeuSerCysValLeuProAspPheLeuGlnGlyArgAla 607

```

```

|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
695 CGCGCTGGCTCAGCTCGTGCCTACCGCATTTCTGCAAGGCGGCGAGCC 744

608 GlyArgTrpValGlyValTrpPheAspGlyLeuLeuHisProAspSerVa 624
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
745 GCGCGCTAGCTGGGGGTCTACTTCTGACGGGCTGCTGACACCGAGACTGT 794

624 tProSerProPheArgValAlaProLeuPheSerLeuProThrGlnLeu 641
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
795 GCCCTCCCGGTTCCGGGCTGCGCCGCTCTTCTCCCTGCGCCGACAGCTGC 844

641 roAlaPheLeuAspAlaLeuGlnGlyCysSerThrSerAlaGlyArg 657
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
845 CGGCTTCTCTGATGCACTGACAGGAGCTGCTCCACTTCCGCGGCGGCA 894

658 ProAlaAspArgValGlnArgValThrGlnAlaLeuArgSerAlaLeuAs 674
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
895 CCGGCGGAGCGGGGTGAGACAGTGCACGAGCGCTGGCGCTCGCCCTGGA 944

674 pSerCysThrSerSerSerGlnAlaProGlyCysCysGlnGlnTrpAspL 691
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
945 CAGCTGTACTTCTAGCTCGGAAGCCCGAGGCTGCTGCGAGAAATGGAGCC 994

691 euGlyProCysThrThrLeuGlu 698
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
995 TGGAGCCCTGCACACTACACTAGAA 1017

seq_name: gb_est1:BB609618

seq_documentation_block:
LOCUS      BB609618      1006 bp      mRNA      linear      EST 26-OCT-2001
DEFINITION BB609618 RIKEN full-length enriched, 18 days embryo Mus musculus
cDNA clone 1110025H02 5', mRNA sequence.
ACCESSION  BB609618
VERSION     BB609618.1 GI:16451334
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus.

REFERENCE
  1 (bases 1 to 1006)
    Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
    Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda
    M., Koye,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
    Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
    D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
    Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
    Muramatsu,M. and Hayashizaki,Y.
    RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
    Unpublished (2001)
    contact: Yoshihide Hayashizaki
    Laboratory for Genome Exploration Research Group, RIKEN Genomic
    Sciences Center(GSC), Yokohama Institute
    The Institute of Physical and Chemical Research (RIKEN)
    1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
    Tel: 81-45-503-9222
    Fax: 81-45-503-9216
    Email: genome-res@sc.riken.go.jp
    URL: http://genome.gsc.riken.go.jp/

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y. and Hayashizaki,Y.

```


TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabp-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1133 row: k column: 20
High quality sequence stop: 697.
Location/Qualifiers

FEATURES
source
1. 697
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:513575"
/lab_host="NCI_CGAP_L19"
/note="Organ: liver; Vector: PCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 141 a 212 c 205 g 139 t
ORIGIN

Alignment_scores:
Quality: 1226.00 Length: 232
Ratio: 5.330 Gaps: 1
Percent Similarity: 99.138 Percent Identity: 98.707

Alignment_block:
US-09-899-471-5 x B1332055

Align seg 1/1 to: B1332055 from: 1 to: 697

```
294 ArgAsnLeuTrpHisIleAlaArgLeuArgValLeuSerProGlyValTr 310
|||||
2 AGGAACTCTGGCAGCATAGCCAGCTGGGCTACTGCTCCCGAGAGTANG 51
|||||
310 pGluLeuAspAlaProCysCysLeuProGlyLysValThrLeuCysTrpG 327
|||||
52 GCAGCTAGATGCGCTTGTCTGTCGCGGCAAGGTACACTGTGCTGCGC 101
|||||
327 InAlaProAspGlnSerProCysGlnProLeuValProProValProGln 343
|||||
102 AGGCACACAGACAGAGTCCCTGCCAGCCACTTGTGCCACAGTCCCGCAG 151
|||||
344 LysAsnAlaThrValAsnGluProGlnAspPhe.GlnLeuValAlaGlyH 360
|||||
152 AAGAGCGCCACTGTGATGAGCCACAAAGATTTCATGTTGGTGGCAGGCC 201
|||||
360 IsProAsnLeuCysValGlnValSerThrTrpGluLysValGlnLeuGln 376
|||||
202 ACCCCCAACCTCTGTCTCAGGTGAGCACTGGGAGAAAGTTCAGCTGCAA 251
|||||
377 AlaCysSerTrpAlaAspSerLeuGlyProPheLysAspAspMetLeuLe 393
|||||
252 GCCTGCTGTGGGGCTGACTCTTGGGGCCCTTCAAGATGATATGCTGTT 301
|||||
393 uValGluMetLysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluP 410
|||||
302 AGTGGAGATGAAAACCGGCTCAACACACATCAGTGTGCTGTGGANAC 351
|||||
410 roSerGlyCysThrProLeuProSerMetAlaSerThrArgAlaAlaArg 426
|||||
352 CCAGTGGCTTACACCACTGCGCCAGCATGCGCTCCACAGAGAGCTCTGCG 401
|||||
427 LeuGlyGluGluLeuLeuGlnAspPheArgSerHisGlnCysMetGlnLe 443
|||||
```

```
402 CTGGAGAGAGACTTGTCTCCAGACTTCGATCACACCAGTGTATGAGCT 451
|||||
443 uTrpAsnAspAspAspMetGlySerLeuTrpAlaCysProMetAspLysT 460
|||||
452 GTGGAGACGATGACACATGAGATGCGCTATGAGCGCTGCCCATGACAGT 501
|||||
460 yTrIleHisArgArgTrpValLeuValTrpLeuAlaCysLeuLeuLeuAla 476
|||||
502 ACATCCACAGCGCGCTGGTCTAGTATGCTGCGCTGCCCTACTCTTGCT 551
|||||
477 AlaAlaLeuPhePhePheLeuLeuLysLysAspArgArgLysAlaAl 493
|||||
552 GCGGCGCTTTCTTCTCTCCCTTCAAAAAGACCGCAGAAAGCGGC 601
|||||
493 AARGLysSerArgTrpAlaLeuLeuLeuHisSerAlaAspGlyAlaGlyT 510
|||||
602 CCGTGGCTCCCGCAGCGCTTGTCTCTCCACTCCGCGAGAGAGCGGCT 651
|||||
510 yTrGluArgLeuValGlyAlaLeuAlaSerAlaLeuSerGlnMet 524
|||||
652 ACGACCGTCTGTGGAGACGACTGCGCTGCTTGAAGCCAGATG 695
|||||
```

seq_name: gb_est2:BG968034
seq_documentation_block:
LOCUS BG968034 784 bp mRNA linear EST 12-JUN-2001
DEFINITION 602832413P1 NCI_CGAP_C024 Mus musculus CDNA IMAGE:4987035 5',
mRNA sequence.
ACCESSION BG968034 GI:14355671
VERSION BG968034.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 784)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabp-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10997 row: g column: 04
High quality sequence stop: 646.
Location/Qualifiers

FEATURES
source
1. 784
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4987035"
/lab_host="NCI_CGAP_C024"
/note="Organ: colon; Vector: PCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 173 a 231 c 229 g 151 t
ORIGIN

alignment_scores:
Quality: 1187.00 Length: 254
Ratio: 5.051 Gaps: 5
Percent Similarity: 92.520 Percent Identity: 89.370

alignment_block:
US-09-899-471-5 x BG968034

430 IuLeuLeuGlnaspPheargSerHisGlnCysMetGlnLeuTyrAsnAsp 446


```

103 AGTTCGTCAGAGACTTCGATCCAGACACAGATGATGACGCTGTGGAACGAT 152
447 ASPASNMTGTLSELEUTRPAACYSPPROMETASPLYSTYRILEHSAR 463
153 GACACAGAGGATCGCTATGGCCCTGCCCCATG. GACAAATGATCATCCACAG 201
463 GATGTTPALLEUVALTRPLEUAIACYSLEUULEUUAIAAIALEU 480
202 GCGCTGGCTCTAGTATGGCTGGCTGCTTACTCTTGGCTGGCGGCTTT 251
480 hepheleuleuleuylslyaspargarplysalaalaaglyser 496
252 TCTTCTTCTCTCTTCTTAAAAAGACCGCAGAGAAACGGCCCTGGCTCC 301
497 ARGTHRALALEUULEUHSERIALAASPGLYALAGLYTYRIGLUARGL 513
302 CGCAGCGGCTTCTCTCTCCACATCCCGCCGAGGAGCGGCTACGAGCGCT 351
513 UVALGVALALEUUAISERIALAASERLIMETPROLEUARYVALAAY 530
352 GGTGGAGACACTGGCGCTCCGCTTGGAGCCAGATGCCATGCCGCTGCCG 401
530 alaspLeuTrpSerArgrgluleuSerAlaHisglYAlaLeuATrp 546
402 TGGACCTGTGAGCGCGCGGAGCTGAGCGCCGACGAGCCCTAGCCTGG 451
547 PHEHISHSGLNARGARGYLLEUENGINGLYGLYVALIILE 563
452 TYACACACACGACGACGCGCTATCCTGAGAGAGGGTGGTGAATTCCT 501
563 uleuPheSerProAlaAlaValAlaGlnCysGlnGlnTrpLeuGlnLeu 580
502 TCTCTCTCTGCGCGCGCGCGGAGGTGTCAGAGTGTGCTGCTGCTGCTCC 551
580 InThrValGluProGlyProHisAspAlaLeuAlaAlaTrpLeuSerCys 596
552 AGACAGTTGAGCCCGCGCGGATGACGCCCTGCGCCCTTGCACACTGC 601
597 VALLEUPROASPHLEUENGLIYARGALATHRGLYARGTYRVALIYVA 613
602 GTGCTTACCGGATTTTCTGCAAGC. CGGGGAGCCGTTAAACCGTGGGGGT 650
613 IYRPHASPGLYLEUENHSIPROASPSERVALPROSERPROPHARY 630
651 CTTACTTGACGGGCGCTGCTGCAACCAACTTGTCTCTCCGCGTTCGGC 700
630 alAlaProLeuPheSerLeuProThrGlnLeuProAlaPheLeuAspAla 646
701 GTGGCCCGGTTTTTCTCCCTGCGGACCTTGCACACTTCTGGAATGA 750
647 leuGlnGlyCysSerThrSer 654
751 CTTAAGGGAAGGTGATCATATTTCG 774

```

seq_name: gb_est2:BF237179

seq_documentation_block:

LOCUS BF237179 722 bp mRNA linear EST 14-NOV-2000
 DEFINITION 602028133F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:4163217 5',
 mRNA sequence.

ACCESSION BF237179

VERSION BF237179.1 GI:11151096

SOURCE

ORGANISM house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 722)
 TITLE NIH-MGC http://mgi.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov

FEATURES

SOURCE

Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.llnl.gov
 Plate: LHAM9447 row: a column: 10
 High quality sequence stop: 584.
 Location/Qualifiers

```

1..722
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4163217"
/clone_id="NCI_CGAP_L19"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Liver; Vector: pCMV-SPOK6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP library."
BASE COUNT
128 a 238 c 255 g 101 t
ORIGIN

```

alignment_scores: Quality: 1058.00 Length: 207
 Ratio: 5.161 Gaps: 0
 Percent Similarity: 99.034 Percent Identity: 97.585

alignment_block:

US-09-899-471-5 x BF237179/rev

Align seg 1/1 to reverse of: BF237179 from: 1 to: 722

```

492 AAlaAargGlySerArgThrAlaLeuLeuHisSerAlaAspGlyAl 508
|||||
629 GCGCGCGCTGGCTCCCGCACAGACCTGCTCTCCGCGCGAGGAGC 580
508 AGLYTYRGLUARGLEUVALGLYALALEUUAISERIALAASERLIMETP 525
|||||
579 GGGCTACAGAGCGCTGTGTGGAGACACTGGCGCTCCGCTGAGCCAGATGC 530
535 IoleuArgValAlaValAspLeuTrpSerArgrgluleuSerAlaHis 541
|||||
529 CACTGCGCTGGCGCGTGGACCTTGGAGCGCGCGAGCTGAGCGCCGAC 480
542 GLYALALEUUALATRPHEHISHSGLNARGARGYLLEUENGINGLY 558
|||||
479 GGAGCCCTAGCTGTGTCACACACAGCGAGCGCGTATCTGCGAGGAGG 430
558 YGLYVALIILEUULEUPHESETRPROAIAAIAAIAAGINCYSGLNG 575
|||||
429 TGGCTGTGTAATCTTCTCTCTGCGCGCGCGGCGCGGACAGTGTAGC 380
575 InTrpLeuGlnLeuGlnTrpValGluProGlyProHisAspAlaLeuAla 591
|||||
379 AGTGCTGACAGCTCCAGACAGTGGAGCCCGGGCGCGCTGACGCCCTGGCC 330
592 AlATrPLeuSerCysValLeuProAspPheLeuGlnGlyArgAlaThrG1 608
329 GCTGGCTACAGCTGCTGTACCCGATTTCTCCAGAGGCGGGGAGACCGG 280
608 YARGTYRVALGLYVALTYRPHASPGLYLEUENHSIPROASPSERVALP 625
|||||
279 CCGCTACCTCGGGGTCTACTTGTGACGGGCTGTGACCCAGACTCTGTGC 230
625 roserProPheArgValAlaProLeuPheSerLeuProThrGlnLeuPro 641
|||||
229 CCTCCCGCTTCGCGCTCCCGCGCTCTTCTCCCTGCTGCGAGCTGCGG 180
642 AlApheLeuAspAlaLeuGlnGlyCysSerThrSerAlaGlyArgPr 658
|||||

```


633 eupheSerleuProthrGluLeuProAlaPheLeuAspAlaLeuGly 649
 243 TCTTCTCCCTGCTGCGACGCTGCGGCTTCTCGATGCTGCGAGGA 194
 650 G1yCSerThrSerAlaGlyArgProAlaAspArgValGluArgValTh 666
 193 GGTGCTCCACTTCCGGGGGGGAGCCCGGAGACGGGTGAGACAGAGTAC 144
 666 rGlnAlaLeuArgSerAlaLeuAspSerCysThrSerSerSerGluAlaP 683
 143 CCAGGCGCTGCGGCTGCGGCTGCGACAGCTGACTTACTGCTCGAAGCCC 94
 683 rG1yCysCysGluGluThrPaspLeuGlyProCysThrThrLeuGlu 698
 93 CAGGCTGCTGCGAGGATGGAGCTGGAGCCCTGCACTACACTAGAA 47

seq_name: gb_est2:BI251117

seq_documentation_block:

LOCUS BI251117 839 bp mRNA linear EST 17-JUL-2001
 DEFINITION 602994315F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5150378 5',
 mRNA sequence.

ACCESSION BI251117

VERSION BI251117.1 GI:14800198

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 839)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM11371 row: m column: 03

High quality sequence stop: 836.

Location/Qualifiers

1. 839

/organism="Mus musculus"

/strain="C57/B6"

/db_xref="taxon:10090"

/clone_image="IMAGE:5150378"

/clone_lib="NCI_CGAP_Mam5"

/tissue_type="tumor, gross tissue"

/dev_stage="7 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: Salt;

Site: 2: NCI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigators

providing samples: Lothar Hennighausen/Robin Humphreys,

NIH"

BASE COUNT 146 a 248 c 268 g 177 t

ORIGIN

alignment_scores:

Quality: 1012.50 Length: 221

Ratio: 4.868 Gaps: 3

Percent Similarity: 94.118 Percent Identity: 93.665

alignment_block:

QS-09-899-471-5 x BI251117

Align seg 1/1 to: BI251117 from: 1 to: 839

1 MetProValSerThrPheLeuSerLeuAlaLeuGlyArgAsnProVal 17
 199 ATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 248
 17 lValValSerLeuGluArgLeuMetGluProGlnAspThrAlaArgCys 34
 249 GGTGCTCTCTGAGAGACTGATGAGCCCTCAGGACACTGACAGCTGCT 298
 34 erLeuGlyLeuSerCysHisLeuThrPaspGlyAspValLeuCysLeuPro 50
 299 CTTGAGCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 348
 51 G1ySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeuG1 67
 349 GGAAGCCCTCAGTCCAGTCCAGGCGCTGCTGCTGCTGCTGCTGCTGCA 398
 67 nThrGluLeuValLeuArgCysProGlnGlyThrAspCysAlaLeuArg 84
 399 GACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 448
 84 AlArgValValValHisLeuAlaValHisGlyHisThrPalaGluProGlu 100
 449 TCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 498
 101 GluAlaGlyLysSerAspSerGluLeuGlnGluSerArgAsnAlaSerle 117
 499 GAAGCTGGAAGTCTGATTCAGAACTCCAGAGAGTCTGGAAGGCTCTCT 548
 117 uGlnAlaGlnValValLeuSerPheGlnAlaTyrProIleAlaArgCys 134
 549 CCAGGCCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598
 134 lAluLeuGlnValGlnValAlProAlaAspLeuValGlnProGlyGlnSer 150
 599 CCTGCTGAGAGTCCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 648
 151 ValGlySerAlaValPheAspCysPheGluAlaSerLeuGlyAlaGluVal 167
 649 GGGGCTTGGCGGTATTTGACTCTTTCGAGGCTAGTCTTGGGCTGAGGT 698
 167 lGlnIleTyrPser.TyrThrLysProArgTyrGlnLysGluLeuAsnLeu 183
 699 ACAGATCTGCTGCTTACACAGAGCCAGGATCCAGAAAGAGCTCAACCTC 748
 184 ThrGlnGlnLeuProAspCysArgGlyLeuGluValArgAspSerIleG1 200
 749 ACACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 797
 200 nSerCysTrpValLeuProThrLeuAsnValSerThrAsp.GlyAspAsn 216
 798 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 820
 217 ValLeuLeu 219
 821 GTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 829
 seq_name: gb_est1:BB637050

seq_documentation_block:

LOCUS BB637050 620 bp mRNA linear EST 26-OCT-2001

DEFINITION BB637050 RIKEN full-length enriched, adult male aorta and vein Mus

musculus cDNA clone A530050M19 5', mRNA sequence.

ACCESSION BB637050 GI:16472835

VERSION BB637050.1

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 620)

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
 Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda

M. Koya, S. Matsuyama, T. Miyazaki, K. Nomura, K. Ohno, M. Ozakazi, Y. Okido, T. Saito, R. Sakai, C. Sakai, K. Sano, H. Sasaki, D. Shibata, K. Shingawa, A. Shiraki, T. Sogabe, Y. Suzuki, H. Tagami, M. Tagawa, A. Takahashi, F. Takeda, Y. Tanaka, T. Toyota, T. Muraatsu, M. and Hayashi, I. Y.

RIKEN Mouse ESTs (Arikawa, T., et al. 2001)

Unpublished (2001)

Contact: Yoshinide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
url:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap trapped selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. *Genome Res.* 10 (10), 1617-1630 (2000)
Wagai,K., Fujiyake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*
10 (11), 1757-1771 (2000)
Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
non-redundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamane,I., Aizawa
K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of full-length mouse cDNAs compared with
human genome sequences. *Mamm. Genome* 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

Location/Qualifiers
1. .620

```

1..620
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A530050M19"
/clone_1fb="RIKEN full-length enriched, adult male aorta
and vein"
/sex="male"
/tissue_type="aorta and vein"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer 15'
GAGAGAGAGAGAGATCCAGACCTCTTTTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalase thermo activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence 15' GAGAGAGAGATCTCCAGCTAATTAATTAATTAATTC
3'). cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
phage I."

```

```

alignment_scores:      Quality: 1007.00      Length: 191
                       Ratio: 5.272      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 99.476

alignment block:
US-09-899-471-5 x BB637050/rev ..

Align seg 1/1 to reverse of: BB637050 from: 1 to: 620

```

Align seg 1/1 to reverse of: BB637050 from: 1 to: 620

```
alignment_block:
    00 000 471 5
```

US-09-899-471-5 x BB637050/rev

```

508 AAGAGTGTGTAAGLeuValAGIValAlaLeuAlaSerAlaLeuSerGlnme 522
518 GCGGGCTACGAGCGTCTGGTGGAGACACTGGCGTCCGCGTTGAGCCAGAT 568
524 tProLeuArgValAlaAlaValAspLeuTrpSerAlaArgValGluLeuSerAlaH 544
541 isglValAlaLeuAlaTrpPheHisHisGlnArgArgArgGluLeuGlnGlu 557
518 ACGGAGCGCTAGCGCTGGTCTCACCACGACGAGACCGCGTATCTCTGCAGAG 468
558 GlyGlyValValIleLeuLeuPheSerProAlaAlaValAlaGlnCysG1 574/4
468 GGTGGCGGTGGTAAACCTCTCTCTCTCGCCGCCGCGCGCTGGCCGAGTGTCA 419
574 nglTrpLeuGlnLeuGlnTrpValGlnProGlyProHisAspAlaLeuAla 599
418 GCGGTGGCTGAGCTCCAGACAGTGGAGCCCGCGGCCGATGACGCCCTCG 368
591 IaAlaTrpLeuSerCysValLeuProAspPheLeuGlnGlyArgAlaThr 607
368 CCGCCGTGGCTCAGCGTGGCGTCTACCCCGATTTCTCGCAAGGCCCGGCGACC 319
608 GLArgTrpTrpAlaGlyValAlaTrpPheAspGlyLeuLeuHisProAspSerVa 622
318 GCGCGTACGTGGGGGTCTACTGCAGCGGGCTGTCTGCACCCAACTCTGT 268
624 lProSerProPheArgValAlaProLeuPheSerLeuProThiGlnLeuP 644
268 GCCCTGCCCGCTTCGGCGTGGCGCCCGCTCTTCCCTGCCCTGCAGCTGC 219
641 roAlaPheLeuAspAlaLeuGlnGlnGlyCysSerThrSerAlaGlyArg 657
218 CGCGTTTCTGTGAGGCACTGTGAGGAGGCGTGTCCACTCCCGCGGGCGCA 168
658 ProAlaAspArgValAGIArgValAlaThrGlnAlaLeuArgSerAlaLeuAs 674
168 CCGCGCGAGCGGGTGGAGAACGACTGACCCAGCGGCTCGGGTCCGCCCTGGA 119
674 pSerCysTrpHisSerSerSerGluAlaProGlyCysCysGlnGluTrpAspL 691
118 CAGCTGTACTTTCAGCTCGGAAAGCCCGCAGGCGTGTCCGAGAAATGGGACC 69
691 euGlyProCysTrpThrLeuGln 698
68 TGGGACCCCTGCACACTACACTAGAA 46

```

seq_documentation_block:

LOCUS	BE290565	562 bp	mRNA	Linear	EST: 26-OCT-2000
DEFINITION	601089427F1	NCI_CGAP_Mam5	Mus musculus	CDNA clone	IMAGE:3484038 5',

Accession: [mRNA sequence](#)
[PF290565](#)

KEYWORDS	EST
VERSION	BE290565.1
REVISION	GI:9171696

SOURCE ORGANISM	house mouse, <i>Mus musculus</i>
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE	1
AUTHORS	(bases 1 to 562) NIH-MGC http://mgc.nci.nih.gov/

BASE COUNT	ORIGIN
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

112 a	199 c	216 g	93 t
-------	-------	-------	------

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Lohar Hemmighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHM8516 row: f column: 07
High quality sequence stop: 562.
Location/Qualifiers
1. 562
/organism="Mus musculus"
/strain="C57/86"
/db_xref="taxon:10090"
/clone_image:3484038"
/clone_lib="NCI-CCAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: Salt;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lohar Hemmighausen/Robin Humphreys,
NIH"

BASE COUNT 86 a 190 c 180 g 106 t

ORIGIN

Alignment_scores:
Quality: 999.00 Length: 187
Ratio: 5.342 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:
US-09-899-471-5 ..

Align seg 1/1 to: BE290565 from: 1 to: 562

```

438 HHSGLNCGYMETGInleutrpnsnaspasnmetyserleutrpai 454
|||||
2 CACCAAGTGTATGCAGCTGTGGAACGATGACACATGGATCGCTATGGGC 51
454 acyspmetasplystfylehisaragrtppvalleuvaltrpleua 471
|||||
52 CTGCCCATGGACAACTACATCCACAGGCGCTGGTCTAGTATGGCTGG 101
471 lacysleuleuleualaalaleupherpetherleuleuleulslys 487
102 CCTGCTCTACTTGTGGCTGGCGCTTTCTTCTCTCTCTCTCTCTCTCT 151
488 asparagatglyalalalargllyseratgthrAlaleuleuleuhisse 504
|||||
152 GACCCGACGAAAGCGGCCGCTGCCCTCCACAGCGCTTGCCTCCACTTC 201
504 talaspalyalaglytyrgluargleuvalaglyalaleuualaseratal 521
|||||
202 CGCGCAGCGAGGCGGCTACGACGCTGTGGTGGAGCAGCTGGCGCTGG 251
521 euserglmetetroleuarqvalalavalaspheutrpseratqarglu 537
252 TGAGCGAATGCTCACTGCGCGTGGCGCTGGACCTGTGGAGCGCCGCGAG 301
538 leuseralalhisglyAlaleualatrrpnehlshsglnahargatrgtl 554
|||||
302 CTGACGCGGACGAGGCCCTAGCTGGTTCACACACAGCAGCGCGGTAT 351
554 eleunglnuglyglyvalalalleuleuleuhereserproalalavalala 571
|||||
351 CCTGACAGAGGCTGCTGTATATCTTCTCTCTCTGCCCCGCGCGCTGG 401

```

```

571 1aglnCysGlnGlnInrPleuGlnLeuGlnIphrValGlnProGlyProHis 587
|||||
402 CGCAGCTGCAGTGCTGCACCTCCAGACAGTGGAGCCCGGCCGCAT 451
588 ASPAlaLeuAlaAlaTrpLeuSerCysValLeuProAspPheLeuGlnG1 604
|||||
452 GAGCGCCCTCGCGCTGCCTCAGCTCGTCTCAGCCATTTCTGCAAG 501
604 yArGAlaThrGlyArGTYrValGlyValIyrrPheAspGlyLeuLeuHisP 621
|||||
502 CCGGCGCAGACCGCGCTACGTCGGGGCTCCTCAGCGGGCTGTCAGCC 551
552 CAGACTCTGTG 562
seq_name: gb|est2:BF160122

seq documentation block:
LOCUS BF160122 635 bp mRNA linear EST 30-OCT-2000
DEFINITION 60167822F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3984149 5',
mRNA sequence.
ACCESSION BF160122
VERSION BF160122
KEYWORDS BF160122.1 GI:11040229
EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 635)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9185 row: 1 column: 06
High quality sequence start: 2
High quality sequence stop: 632.
Location/Qualifiers
1. 635
/organism="Mus musculus"
/strain="C2BECH II (fetal)"
/db_xref="taxon:10090"
/clone_image="3984149"
/clone_id="NCI_CGAP_Lu29"
/tissue_type="spontaneous,tumor, metastatic to mammary.
stem cell origin."
/lab_host="DH10b"
/notes="Organ: lung; Vector: pGEM-SPORT6; Site: 1; Salt;
Site: 2; Notti; Cloned using directionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 92 a 229 c 204 g 110 t
ORIGIN

alignment_scores:
Quality: 975.00 Length: 207
Ratio: 4.924 Gaps: 4
Percent Similarity: 95.652 Percent Identity: 93.237

alignment block:
US-09-899-471-5 x BF160122 ..

Align seg 1/1 to: BF160122 from: 1 to: 635

```



```

494 ArgGlySerArgThrAlaLeuLeuHisSerAlaAspGlyAlaGlyTyr 510
      |||.....|||
2 CGTGGCGCCGACGCGCTTGTCTCCATCCGCGCCGACGAGCGGCTA 51
510 rGluArgLeuValGlyAlaLeuAlaSerAlaLeuSerGlnMetProLeu 527
      |||.....|||
52 CGAGGCTGTGTAGAGACACTGGCTGCGGCTGTGACCAATGCGCACTGC 101
527 rGValAlaValAspLeuTrpSerArgArgGluLeuSerAlaHisGlyAla 543
      |||.....|||
102 GCGTGGCGCTGTGACCTGTGAGCGCGCGAGCTGACGCGGACGAGACC 151
544 LeuAlaTrpPheHisHisGlnArgArgGlyLeuGlnGlnGlyAla 560
      |||.....|||
152 CTAGCGTGTGTCCACCAACAGCGCGCGCTATCTCGACGAGATGGCGT 201
560 lValIleLeuLeuPheSerProAlaAlaValAlaGlnGlnGlnTrpL 577
      |||.....|||
202 GGTATCTCTTCTCTCTCCCGCGCGCGCTGTGCGACGTGTGACAGTGGC 251
577 euglLeuGlnThrValGluProGlyProHisAspAlaLeuAlaAlaTrp 593
252 TGCAAGCTCCAGACAGTGAAGCCGCGTCCGATGACGCGCTCGCGCTCG 301
594 LeuSerCysValLeuProAspPheLeuGlnGlyArgAlaThrGlyArgTyr 610
302 CTCAGCTCGTGTCTACCCGATTTCTGCAAGCGCGCGCGCGCTCA 351
610 rValGlyValTyrPheAspGlyLeuLeuHisProAspSerValProSerP 627
352 CGTGGCGCTGTCTCTCTCCCGCGCGCGCTGTGACCACTGTGTGCTGCC 401
627 rPheArgValAlaProLeuPheSerLeuProThrGlnLeuProAlaPhe 643
402 CGTCCGCGTGGCGCGCGCTCTCTCTCCGCTCGCAGTGGCGGCTTTC 451
644 LeuAspAlaLeuGlnGlyCysSerThrSerAlaGlyArgProAlaA 660
452 CTGAGATGACATCGACAGGAGCTGCTCCACGTCGCCGCGGCGACCGCA 500
660 sPArgValAlaArgValThrGlnAlaLeuArgSerAlaLeuAspSerC 676
501 ..CGGGTGAACAGATGGCCAGCGCTGC...AGTCCGCTGTGACACACTG 545
676 sThrSerSerSerGlnAlaProGlyCysGlnGluGluTrpAspLeuGlyP 693
546 TACTTCTAGCTCGGAAGCCCGAGGCTGTGCGGAGATGGAGACTGGGAC 595
693 rGlySerThrThrLeuGlu 698
      |||.....|||
596 CCTGCACATACACTAGAA 612

```

seq_name: gb_est2:BG761977

seq_documentation_block:

LOCUS BG761977 935 bp mRNA linear EST 15-MAY-2001
DEFINITION 602718928F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4858806 5',
mRNA sequence.

ACCESSION BG761977
VERSION BG761977.1 GI:14072630

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 935)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.femail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LUNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LUNL at:
http://image.llnl.gov
Plate: LCM1713 row: h column: 07
High quality sequence stop: 813.
Location/Qualifiers
1. 935
/organism="Homo sapiens"
/ab_xref="taxon:9606"
/clone="IMAGE:4858806"
/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site: 2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH-MGC
Library."

BASE COUNT 154 a 301 c 291 g 189 t
ORIGIN

alignment_scores:
Quality: 912.50 Length: 307
Ratio: 3.755 Gaps: 8
Percent Similarity: 79.153 Percent Identity: 62.866

alignment_block:
US-09-899-471-5 x BG761977 ...

Align seg 1/1 to: BG761977 from: 1 to: 935

```

261 ThrAspLeuValProCysLeuGlnGlnValTrpSerLeuGluProAs 277
      |||.....|||
2 ACAGACCTGCTTCCCTGCTGTGATTCAGGTGCGCTGTGGAACCTGA 51
277 pSerGluArgValGluPheCysProPheArgGluAspProGlyAlaHis 293
      |||.....|||
52 CTCGCTTGAAGACACATCTGCCCTTCAGGGAGAGACCCCGCCACAC 101
294 ArgAspLeuTrpHisIleAlaArgLeuArgValLeuSerProGlyValTr 310
      |||.....|||
102 CAGAACCTCTGGCAGAGCGCGCGCTGCTGACCTGACCCGACAGAGTG 151
310 pGlnLeuAspAlaProCysCysLeuProGlyLysValThrLeuGlyTrp 327
152 GCTGCTGAGACGACCGTGTCTGCTGCGCCGACGAGAGGCGACTGTCTGC 201
327 lAlaProAspGlnSerProCysGlnProLeuValProProValProGln 343
202 GGGCTCCGGGTGGGAGACCTGCCACGCACTGGGAGGCACTTCTCTGG 251
344 LysAspAlaThrValAsnGluProGlnAspPheGlnLeuValAlaGly 360
252 GAGAAATGTACTGTG...GACGTTCTGAGTTCATTCGGAAGGCA 298
360 sProAsnLeuCysValGlnValSerThrTrpGluLysValGlnLeuGln 377
299 CCTTAACCTCTGTGTTCAGTGAACACTGCGAGAACCTGACGTCGACG 348
377 lAcSerTrpAlaAspSerLeuGlnProPheLysAspAspMetLeuLeu 393
349 AGTGTCTTGTGGGTGACTCCTGTGGGCTCTCTCAAGAGGAGTGTACTG 398
394 ValGluMetLysThrGlyLeuAsnAsnThrSerValCysAlaLeuGluPr 410
      |||.....|||

```



```

/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:494084"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10b"
/notes="Organ: mammary; Vector: pcwv-SPOR6; Site: 1: Saliv. site; 2: NotI. Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      128 a      229 c      259 g      171 t
ORIGIN

alignment_scores:
      Quality: 908.00      Length: 213
      Ratio: 4.563      Gaps: 8
Percent Similarity: 93.427      Percent Identity: 91.080

alignment_block:
US-09-899-471-5 x BG917539 ..

Align seg 1/1 to: BG917539 from: 1 to: 787

1 MetProValSerTrpPheLeuLeuSerLeuAlaLeuGlyArgAsnProVa 17
|||||
163 ATCCGCTGCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 212
|||||
17 ValValSerLeuGluArgLeuMetGluProGlnAspThrAlaArgCys 34
|||||
213 GGCCTCTCTCTCTGAGAGACTGATGAGACCTTACAGACACTGCACGCTG 262
|||||
34 LeuGluLeuSerCysHisLeuTrpAspGlyAspValLeuCysLeuPro 50
|||||
263 CTCTAGAGCCCTCTCTCTGCCACCTCTGGAGATGGAGAGTGTCTGCTCT 312
|||||
51 GlySerLeuGlnSerAlaProGlyProValLeuValProThrArgLeu 67
|||||
313 GGAAGCTCCAGTCTGCCCAAGGCCCTGTGCTAGTGCCTACCCGCGTGA 362
|||||
67 nThrGluLeuValLeuArg.CysProGlnLysThrAspCysAlaLeu 83
|||||
363 GAGCGAGCTGTGCTGAGCTGTTCACAGAACAGATGGCCCTCTGT 412
|||||
84 ValArgValVal.ValHisLeu.AlalaHisGlyHisTrp.AlaGlu 99
|||||
413 GTCCCGTGTGTGTGCTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTG 462
|||||
99 OGUGluAlaGlySerAspSerGluLeuGlnGlnLysSerArgAsnAla 116
|||||
463 TGAAGAAGCTGGAAGAAGTGTGATTGAGAACTCCAGAGATGAGAAC 512
|||||
116 LeuLeuAlaGlnValValLeuSerPheGlnAlaTrpProIleAlaArg 132
|||||
513 CTCTCCAGGCCAGGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 562
|||||
133 CysAlaLeuLeuGluValGlnValProAla.AspLeuValGlnProGly 149
|||||
563 TGTGCTGCTGTGAGGTCCAGGTCCCGCTGACCTGAGGAGAGCCGTG 612
|||||
149 InsTrpAlaGlySerAlaValPheAsp.CysPheGluAlaSer.LeuGly 165
|||||
613 AGTCCGTGCTGTGCGGTATTGACTGTGTCGAGGCTAGTCTTGCGGG 662
|||||
165 Laglu.ValGlnIleTrpSerTrpThrLysProArgTyrGlnLysGlu 181
|||||
663 CTGAGGCTACAGATCTGTCTCTACACGAAGCCAGGGATACCAAGAAAG 712
|||||
181 uAsnLeuThrGlnGlnLeuProAspCysArgGlyLeuGlnValArgAsp 198
|||||
713 TCATACACACACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 762

```


198 erlleglnsercysrtpvalleu 205
 |||||
 763 GCATCCAGAGCTGCTGGATGCTG 785

seq_name: gb_est2:B1765330

seq_documentation_block:

LOCUS B1765330 922 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603050641F1 NIH_MGC_116 Homo sapiens CDNA clone IMAGE:5190442 5',
 mRNA sequence.

ACCESSION B1765330
 VERSION B1765330.1 GI:15756908
 KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 922)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@b-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM11476 row: b column: 11
 High quality sequence stop: 794.
 Location/Qualifiers

FEATURES

source

1. 922
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5190442"
 /clone_lib="NIH_MGC_116"
 /lab_host="DH10B"
 /note="Organ: pooled colon, kidney, stomach; Vector:
 pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 26 yo male, 49 yo
 female, 71 yo male colon; 46 yo male kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH_MGC Library."
 BASE COUNT 147 a 297 c 272 g 206 t
 ORIGIN

alignment_scores:

Quality: 902.00 Length: 344
 Ratio: 3.667 Gaps: 7
 Percent Similarity: 71.512 Percent Identity: 56.977

alignment_block:

us-09-899-471-5 x B1765330 ..

Align seg 1/1 to: B1765330 from: 1 to: 922

27 ProGlnAspThrAlaArgCysSerLeuGlyLeuSerCysHisLeuTPas 43
 |||||
 6 CCTCAGGAGGCTACCCAGCTGCTCCGGGCTCTCCGCGCCCTGCGGA 55
 43 pGlyAspValLeuProGlySerLeuGlnSerAlaProGlyPro 60
 |||||
 56 CAGTACATACACTCTGCTGCTGGGACATGCTGCTGCTGCTGCTGCTG 105

60 aLeuValProThrArgLeuGlnThrGluLeuValLeuArgCysProGln 76
 |||||
 106 TGCTGGCGCCTACGACACCTGACACAGAGCTGCTGCTGCTGCTGCTG 155
 77 LysThrAspCysAlaLeuArgValAlaValAlaHisLeuAlaValH 93
 |||||
 156 GAGACCGAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 205
 93 sGlyHisTrpAlaGluProGluGluAlaGlyLysSerAspSerGluLeu 110
 |||||
 206 T..... 206
 110 InGlnSerArgAsnAlaSerLeuGlnAlaGlnAlaValLeuSerPheGln 126
 |||||
 207GCCTCTCCAGGCCCCAATCGCTGCTCTCTCCAG 242
 127 AlaTrpProIleAlaArgCysAlaLeuLeuGluValGlnAlaProAla 143
 |||||
 243 GCCTACCTCTACTGCCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 292
 143 PLeuValGlnProGlyGlnSerValGlySerAlaValAlaPheAspCysPhe 160
 |||||
 293 CTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 342
 160 LuAlaSerLeuGlnAlaGluValGlnIleTrpSerTrpThrLysProArg 176
 |||||
 343 AGCGTCCGCTAGGAGTGAGTGAAGATGCTGCTGCTGCTGCTGCTGCTG 392
 177 TyrGlnLysGluLeuAsnLeuThrGlnGlnLeuProAspCysArgLys 193
 |||||
 393 TACGAGAGGAGGACTCAACACACAGCAGCTGCTGCTGCTGCTGCTGCT 433
 193 uGluValArgAspSerIleGlnSerCysTrpValLeuProTrpLeuAsn 210
 |||||
 434GCCTGCTCAACG 446
 210 aLserThrAspGlyAspAsnValLeuLeuThrLeuAspValSerGln 226
 |||||
 447 TGTCAGCAGATGCTACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 496
 227 GlnAspPheSerPheLeuLeuThrLeuArgProValProAspAlaLeu 243
 |||||
 497 CAGACCTTGGCGCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 520
 243 sSerLeuTrpTrpLysAsnLeuThrGlyProGlnAsnIleThrLeuAsn 260
 |||||
 521TGC.....AATCAGACTGACCGCAGATCAATACCTGAAC 557
 260 IsthAspLeuValProCysLeuGlnIleGlnValTrpSerLeuLys 276
 |||||
 558 ACACAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 607
 276 oAspSerGluArgValGluPheCysProPheArgGlnAspProGlyAla 293
 |||||
 608 TGACTCGCTTAGGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 657
 293 IArgAsnLeuTrpHisIleAlaArgLeuArgValGlnSerProGlyVal 309
 |||||
 658 ACCAGAACCTCTGCGAACG.CGCCGACTGCTGCTGCTGCTGCTGCTG 706
 310 TrpGlnLeuAspAlaProCysLeuProGlyLysValThrLeuLysTr 326
 |||||
 707 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 756
 326 pGlnAlaProAspGlnSerProCysGlnProLeuValProProValPro 343
 |||||
 757 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 800
 343 LnyLysAsnAlaThrValAsnGluProGlnAsp...PheGlnLeuAla 358
 |||||
 801 ACCTTGAGAGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 850
 359 GlyHisProAsnLeuCysValGlnValSer 368

|||||
851 GGGACACCTACCTCTGTGCTGCATGAC 880

seq_name: gb_est2:B1282622

seq_documentation_block:

LOCUS B1282622 619 bp mRNA linear EST 19-JUN-2001
DEFINITION UI-R-CW05-ccd-b-11-0-UI.s1 UI-R-CW05 Rattus norvegicus cDNA clone

ACCESSION UI-R-CW05-ccd-b-11-0-UI.3, mRNA sequence.

VERSION B1282622

KEYWORDS B1282622.1 GI:14933575

SOURCE EST.

ORGANISM Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 619)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL 9704447

MEDLINE

COMMENT

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@iue.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
non-normalized rat aorta pool library cDNA library preparation:
M.B. Soares Lab Clone distribution: clones will be available
through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA-Yes.

FEATURES

source

1. 619

/location/Qualifiers

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-CW05-ccd-b-11-0-UI"

/clone_11b="UI-R-CW05"

/dev_stage="ADULT"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73b-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CW05

library is a non-normalized library constructed from the

following rat aorta tissues: embryonic day 19, embryonic

day 21, adult day 1, adult day 12, adult day 75, adult day

200. For a detailed description of the library from which

this clone was derived, please visit our web site at

ratseq.eng.uiowa.edu. The subtraction has been previously

described in (Bonaldo, Lennon and Soares, Genome Research

6:791-806, 1996)

TAG_LIB=UI-R-CW05

TAG_TISSUE=rat aorta pool

BASE COUNT 105 a 197 c 211 g 106 t

ORIGIN

alignment_scores:

Quality: 893.00

Ratio: 4.907

Percent Similarity: 94.301

Length: 193

Gaps: 0

Percent Identity: 89.119

Alignment block:

US-09-899-471-5 x B1282622/rev

Align seg 1/1 to reverse of: B1282622 from: 1 to: 619

506 AspG1yAlaGlyTrpGluArgLeuValGlyAlaLeuAlaSerAlaLeu 522
|||||
619 GAGGAGACATCTACGAGCGCTGTGGGGCTGTGGGCTGCGCTTACG 570
|||||
522 rGlnMetProLeuArgValAlaValaAspLeuTrpSerArgArgGluLeu 539
|||||
569 CCAGATCCACTGCGCGCTGCGGAGCTGTGGAGCTGTGGAGCCGCGAGCTGA 520
|||||
539 erAlaHisGlyAlaLeuAlaTrpPheHisHisGlnArgArgGluLeu 555
|||||
519 CCGCGACGAGCCCTGCGCTGTGCTCCACGACACGCGCGCGGACCTG 470
|||||
556 GlnGluGlyValValAlaLeuLeuPheSerProAlaAlaValAlaG1 572
|||||
469 CAGGAGATGGCGGTGCTACTTCTTCCACCGCGCGCGCTGCGCA 420
|||||
572 nCysGlnGlnTrpLeuGlnLeuGlnTrpValGluProGlyProHisAsp 589
|||||
419 GTGTACACAGTGGCTGCGAGCTCCAGACGCTGGAGCCGCGCGCATGACG 370
|||||
589 laLeuAlaAlaTrpLeuSerCysValLeuProAspPheLeuGlnArg 605
|||||
369 CTCTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 320
|||||
606 AlaThrGlyArgTrpValGlyValTrpPheAspGlyLeuLeuHisProAs 622
|||||
319 GCGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 270
|||||
622 pSerValProSerProPheArgValAlaProLeuPheSerLeuProTrpG 639
|||||
269 CACTGTACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 220
|||||
639 lneuproAlaPheLeuAspAlaLeuGlnGlyGlyCysSerThrSerAla 655
|||||
219 AGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 170
|||||
656 G1yArgProAlaAspArgValGluArgValThrGlnAlaLeuArgSerAl 672
|||||
169 GGGCGACCTGCGGACGCGGAGCGGAGAGCTGCGCGCTGCTGCTGCGC 120
|||||
672 alaAspSerCysThrSerSerSerGluAlaProGlyCysCysGlnGluT 689
|||||
119 CCTAGAGAGCTGTACTTAGAGCTCGAAGCCCGAGGCTGCGCGAGGAT 70
|||||
689 rpaSpLeuGlyProCysThrThrLeuGlu 698
|||||
69 GGGACCTGGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 41